Entwicklung real-time-PCR-basierter Methoden für die moderne DNA-Analytik

Entwicklung real-time-PCR-basierter Methoden für die moderne DNA-Analytik

Inauguraldissertation

zur

Erlangung des Grades

Doktor der Naturwissenschaften

- Dr. rer. nat. -

des Fachbereichs Biologie, Chemie und Geowissenschaften, FB 08

der Justus-Liebig-Universität Gießen

vorgelegt von

Dipl.-Biol. Jochen Wilhelm

geboren am 08.02.1972 in Gießen

Gießen, 2003

Die vorliegende Arbeit wurde im Rahmen des Graduiertenkollegs "Molekulare Biologie und Pharmakologie" am Institut für Biochemie (FB 08) der Justus-Liebig-Universität Gießen in der Zeit vom 01.03.1998 bis 24.12.2002 unter Leitung von Prof. Dr. A. M. PINGOUD angefertigt.

Dekan: PROF. DR. JÜRGEN JANEK
Institut für physikalische Chemie, FB 08
der Justus-Liebig-Universität Gießen
Heinrich-Buff-Ring 58, 35392 Gießen

Erstgutachter: PROF. DR. ALFERD M. PINGOUD
Institut für Biochemie, FB 08
der Justus-Liebig-Universität Gießen
Heinrich-Buff-Ring 58, 35392 Gießen

Zweitgutachter: PROF. DR. MANFRED KRÖGER
Institut, FB 08
der Justus-Liebig-Universität Gießen
Heinrich-Buff-Ring 58, 35392 Gießen

Tag der Promotion: 14.03.2003

A. VERZEICHNIS WISSENSCHAFTLICHER PUBLIKATIONEN

1. Tews, B., Wilhelm, J., Summerer, D., Marx, A., Friedhoff, P., Pingoud, A., Hahn, M. (2003) Application of primers with C4´ alkylated deoxyribose in allele specific Real-time-PCR for increased selectivity in chimerism analysis. *Nucleic Acids Res*, zur Veröffentlichung eingereicht.

2. Wilhelm, J., Pingoud, A., Hahn, M. (2003) Real-time competitive asymmetric PCR with one-color two-point detection for the high-precision quantifications of nucleic acids. *Nucleic Acids Res*, zur Veröffentlichung eingereicht.

3. Wilhelm, J., Pingoud, A., Hahn, M. (2003) Real-time PCR based method for the estimation of genome sizes. *Nucleic Acids Res*, zur Veröffentlichung eingereicht.

4. Wilhelm, J., Pingoud, A., Hahn, M. (2003) Validation of an algorithm for automatic quantification by real-time PCR. *Analytical Biochemistry*, zur Veröffentlichung eingereicht.

5. Al-Fakhri, N., Wilhelm, J., Hahn, M., Heidt, M., Hehrlein, F.W., Endisch, A.M., Hupp, T.M., Cherian, S., Bobryshev, Y.V., Lord, R.S.A., Katz, N. (2003) Simultaneous upregulation of disintegrin metalloproteinases ADAMs 9 and 15 with integrins $\alpha 5\beta 1$ and $\alpha v\beta 3$ in human arteriosclerosis compared with normal arteries. *J Cell Biochem*, im Druck.

6. Wilhelm, J., Pingoud, A., Hahn, M. (2003) SoFAR: Software for fully automatic and highly accurate evaluation of real-time PCR data. *BioTechniques* **34**: 324-332.

7. Königshoff, M., Wilhelm, J., Bohle, R.M., Pingoud, A., Hahn, M. (2003) HER-2/neu gene copy number quantified by real-time PCR: Comparison of gene amplification, heterozygosity, and immunohistochemical status in breast cancer tissue. *Clinical Chemistry* **49**: 219-229.

8. Wilhelm, J., Reuter, H., Tews, B., Pingoud, A., Hahn, M. (2002) Detection and quantification of insertion/deletion-variations by allele-specific real-time PCR: application for genotyping and chimerism analysis. *Biological Chemistry* **383:** 1423-1433.

9. Wilhelm, J. (2002) Akkumulation von Schäden im Erbgut. In: Becker, S.E., Birringer, J., Friedl, C.H., Kollas, A.-K., Nedvetsky, P., Volk, U., Wilhelm, J., Zahner, D. (Hrsg.): *Warum Lebewesen altern – Theorien und Modelle aus Biologie und Medizin*. Logos-Verlag, Berlin, 7-25.

10. Hahn, M., Königshoff, M., Wilhelm, J. (2001) LightCycler Real-time-PCR zur präzisen Quantifizierung der Genom und Transkriptkopienanzahl tumorrelevanter Gene. *BIOspektrum* **2**: 118-121.

11. Hahn, M., Wilhelm, J., Pingoud, A. (2001) Influence of fluorophor dye labels on the migration behaviour of polymerase chain reaction-amplified short tandem repeats during denaturing capillary electrophoresis. *Electrophoresis* **22**: 2691-2700.

12. Wilhelm, J., Pingoud, A., Hahn, M. (2000) A comparison between *Taq* DNA polymerase and its Stoffel fragment for quantitative real-time PCR with hybridization probes. *BioTechniques* **30**: 1052-1062.

13. Wilhelm, J., Hahn, M., Pingoud, A. (2000) Influence of target DNA melting behavior on real-time PCR quantification. *Clinical Chemistry* **46**: 1738-1743.

14. Wilhelm, J., Pingoud, A., Hahn, M. (2000) Detection of *p53* allele deletion in human cancer by quantification of genomic copy number. In: Meuer, S., Wittwer, C., Nakagawara, K. (Eds.): *Rapid Cycle Real-Time PCR. Methods and Applications*, Springer-Verlag, Heidelberg: 159-170.

15. HAHN, M. & WILHELM, J. (2000) LightCycler Real-time-PCR zur präzisen Quantifizierung der Genkopienanzahl tumorrelevanter Gene. In: *PCR – Methoden und Anwendungen*, Köhler-Verlag, Gießen: 101-109.

16. HAHN, M., THÖMMES, C., WILHELM, J., MICHEL, J. (2000) Quantification of retrotransposon XIR-2.5 copy number in genomes of Poeciliidae species. In: MEUER, S., WITTWER, C., NAKAGAWARA, K. (Eds.): *Rapid Cycle Real-Time PCR. Methods and Applications*, Springer-Verlag, Heidelberg: 399-408.

17. WILHELM, J., HAHN, M., PINGOUD, A. (1999) LightCycler based real-time PCR for quantification of *p53* and *p16* tumour suppressor gene copy numbers in bladder tumours. *Biological Chemistry* **380**: S153.

B. VORTRÄGE

1. WILHELM, J. (2002) LC_{32}: Data analysis for real-time quantitative PCR. *CeBIT 2002*, Forschungsrepräsentation der Justus-Liebig-Universität Gießen im TechnologieTransferNetzwerk Hessen auf dem Gemeinschaftsstand der hessischen Hochschulen, Hannover, 13.-20.3.2002.

2. WILHELM, J., HAHN, M., PINGOUD, A. (2001) Die Polymerase-Kettenreaktion: Eine revolutionäre Technik – nicht nur für die Forschung. Techno Now 2002, Forschungsrepräsentation des Landes Hessen, Limburg, 8.3.2002.

3. WILHELM, J., HAHN, M., PINGOUD, A. (2001) Real-time PCR detection and quantification of SNPs and other genetic variations. *Herbsttagung der Gesellschaft für Biochemie und Molekularbiologie*, Bochum, 9.-12.9.2001.

4. WILHELM, J. (2001) Automatische und sichere Auswertung von real-time Quantifizierungen für die molekulare DNA-Analytik. *TransMIT-Seminar „Medizin der Zukunft – Wettbewerbsvorteil durch innovative Erfindungen"*, Gießen, 9.5.2001.

5. WILHELM, J., PINGOUD, A., HAHN, M. (1999) Simple and rapid determination of genome sizes using the LightCycler technique. *LightCycler-Symposium und Anwender-Workshop*, Justus-Liebig-Universität Gießen, Gießen, 9.-10.12.1999.

6. WILHELM, J., HAHN, M., PINGOUD, A. (1998) Quantification of DNA using real-time PCR detection based on fluorescence resonance energy transfer: first results with the LightCycler. *Workshop „Bioanalytik für junge Wissenschaftler" der Gesellschaft für Biochemie und Molekularbiologie*, Schloss Rauischholzhausen bei Marburg, 11.-13.11.1998.

C. POSTER- UND TAGUNGSBEITRÄGE

1. WILHELM, J. (2002) LC32: Data analysis for real-time quantitative PCR. *CeBIT 2002*, Forschungsrepräsentation der Justus-Liebig-Universität Gießen im TechnologieTransferNetzwerk Hessen auf dem Gemeinschaftsstand der hesseischen Hochschulen, Hannover, 13.-20.3.2002.

2. KUHNERT, S., WILHELM, J., HINSCH, E., HINSCH, K.-D. (2002) Set up of a method for the quantitative determination of ZP1, ZP2, and ZP3 mRNA in marmoset monkey oocytes. *35. Annual Meeting on Physiology and Pathology of Reproduction and 27. Veterinary-Human Medical Communal Meeting*, Leipzig, 14.-15.2.2002.

3. Wilhelm, J., Hahn, M., Pingoud, A. (2001) Real-time PCR detection and quantification of SNPs and other genetic Variations. *Herbsttagung der Gesellschaft für Biochemie und Molekularbiologie, Bochum*, 9.-12.9.2001.

4. Hahn, M., Wilhelm, J., Pingoud, A. (2001) Fluorophor specific retardation of PCR amplified STRs during denaturing capillary electrophoresis. *Herbsttagung der Gesellschaft für Biochemie und Molekularbiologie*, Bochum, 9.-12.9.2001.

5. Wilhelm, J., Hahn, M., Pingoud, A. (2001) Improvements in quantitative real-time PCR. *27th Meeting of the Federation of European Biochemical Societies*, Lissabon, 30.6.-5.7.2001.

6. Wilhelm, J., Pingoud, A., Hahn, M. (2000) LightCycler based real-time PCR for high accuracy quantification of gene copy numbers. *18th International Congress of Biochemistry and Molecular Biology "Beyond the Genome"*, Birmingham, 16.-20.7.2000.

7. Hahn, M., Thömmes, C., Wilhelm, J., Michel, J. (1999) Quantification of retrotransposon XIR-2.5 copy number in genomes of Poeciliidae species. *LightCycler-Symposium und Anwenderworkshop der Justus-Liebig-Universität Gießen und Roche Diagnostics*, Gießen, 9.-10.12.1999.

8. Freist, A., Wilhelm, J., Hahn, M., Pingoud, A., Tabiti, K., Seibl, R. (1999) Schneller und exakter Nachweis von Krebszellen in Leukämie-Patienten für Therapie und Prognose. *Tag der offenen Tür des Verbands der chemischen Industrie in Kooperation mit der Justus-Liebig-Universität Gießen*, Gießen, 25.10.1999.

9. Hahn, M., Wilhelm, J., Pingoud, A. (1999) LightCycler based real-time PCR for quantification of *p53* and *p16* tumour suppressor gene copy numbers in bladder tumours. *Herbsttagung der Gesellschaft für Biochemie und Molekularbiologie*, Hamburg, 5.-8.10.1999.

10. Wilhelm, J., Pingoud, A., Hahn, M. (1999) Algorithm for standardized data evaluation of real-time quantitative PCR experiments. *Gemeinsame Jahrestagung der Deutschen Gesellschaft für Klinische Chemie (DGKC) und der Deutschen Gesellschaft für Laboratoriumsmedizin (DGLM)*, Regensburg, 3.-5.10.1999.

11. Wilhelm, J., Hahn, M., Pingoud, A. (1999) LightCycler based real-time PCR for quantification of *p53* and *p16* tumour suppressor gene copy numbers. *5. Werkstattberichte aus Natur- und Lebenswissenschaften der Justus-Liebig-Universität Gießen*, Gießen 9.-10.7.1999.

12. Hahn, M., Wilhelm, J., Pingoud, A. (1999) Quantification of DNA using real-time PCR detection based on fluorescence resonance energy transfer: First results with the LightCycler for human *p53* tumour suppressor gene. *Gene Quantification Europe: Clinical Applications and Drug Development*, München, 13.-14.5.1999.

13. Wilhelm, J., Hahn, M., Pingoud, A. (1999) Quantification of DNA using real-time PCR detection based on fluorescence resonance energy transfer: first results with the LightCycler. *Workshop „Bioanalytik für junge Wissenschaftler" der Gesellschaft für Biochemie und Molekularbiologie*, Schloss Rauischholzhausen bei Marburg, 11.-13.11.1998.

D. Patente

1. Marx, A., Tews, B., Wilhelm, J., Friedhoff, P., Hahn, M., Pingoud, A. (2002) Verbessertes Verfahren zur allelspezifischen PCR. *Europäische Patentanmeldung* Nr. 21139 EP-HIL.

E. GRUNDLEGENDE DEFINITIONEN

- Real-time-PCR (RT-PCR) bezeichnet ein PCR-Verfahren, bei dem während der PCR die Menge der entstandenen Produkte in Echtzeit gemessen wird. Der englische Begriff, in der Literatur manchmal auch als kinetische PCR (*kinetic PCR*) bezeichnet, lässt sich daher mit Echtzeit-PCR übersetzen. Da der aus dem Englischen stammende Begriff RT-PCR inzwischen jedoch auch im Deutschen üblich geworden ist, wird er in dieser Arbeit durchgehend verwendet.

- Ebenso aus dem Englischen übernommen wurden die in diesem Feld häufig verwendeten Begriffe Primer, welcher ein kurzes in der PCR benötigtes Oligodesoxyribonukleotid beschreibt, und Peak, der eine Spitze bzw. einen Durchgang durch ein Maximum einer Funktion oder Messwertreihe bezeichnet.

- Die in RT-PCR-Experimenten gemessene Größe ist die Fluoreszenzintensität bei einer oder mehreren Wellenlängen. Die Angaben der Fluoreszenzintensitäten in Diagrammen verstehen sich immer als relative Fluoreszenzeinheiten. Angaben von Intensitätsverhältnissen (z. B. Akzeptor/Donor) sind entsprechend dimensionslos.

- Die Nomenklatur vieler Parameter in der Auswertung von RT-PCR-Messungen in der Fachliteratur ist uneinheitlich. Bezeichnungen wie *noise band* und *threshold* sowie *crossing point* und *threshold cycle* werden häufig synonym verwendet. In dieser Arbeit werden die Begriffe *noise band* und *threshold* als Schwellenwert bezeichnet, *crossing point* und *threshold cycle* als Schwellenwertzyklus oder kurz als C_T-Wert.

- Als Dezimaltrennzeichen wird in dieser Arbeit der Dezimalpunkt verwendet.

- Zur besseren Lesbarkeit sind Parameter und Funktionen mathematischer Formulierungen in der Schriftart Times angegeben, wobei die Parameter kursiv geschrieben sind. Aus gleichem Grund sind alle die Software betreffenden Bezeichnungen und Ausdrücke in der Schriftart Courier wiedergegeben.

F. LEBENSLAUF

PERSÖNLICHE DATEN

Name	Wilhelm
Vorname	Jochen
Wohnort	Hauptstraße 5, Ebersgöns 35510 Butzbach
Geburtstag	08.02.1972
Geburtsort	Gießen (Hessen)

SCHULAUSBILDUNG

1978 – 1982	Grundschule MPS Oberer-Hüttenberg
1982 – 1991	Gymnasium Weidigschule Butzbach
Juni 1991	Abschluss mit Abitur Hauptfächer Biologie und Mathematik Nebenfächer Deutsch und Religion

WEHRDIENST

1991 – 1992	Wehrdienst in Hermeskeil und Kassel

HOCHSCHULSTUDIUM

Oktober 1993	Beginn des Diplom-Studiums der Biologie im Fachbereich Biologie der Justus-Liebig-Universität Gießen
April 1995	Diplom-Vorprüfung in Biologie Hauptfächer Botanik und Zoologie Nebenfächer Chemie und Physik
Oktober 1995 bis April 1998	Teilnahme am Nebenstudiengang Praktische Informatik im Fachbereich Mathematik der Justus-Liebig-Universität Gießen
Februar/März 1997	Diplom-Hauptprüfung in Biologie Hauptfächer Botanik und Zoologie Nebenfächer Tierphysiologie und Zellbiologie
April 1997 bis Dezember 1997	Diplomarbeit im Institut für Tierphysiologie der Justus-Liebig-Universität Gießen unter Anleitung von Prof. Dr. F. Jauker zum Thema „Entwicklung eines Programms zur Datenakquisition fluorimetrischer Calciummessungen: FluoroDAS"
Dezember 1997	Abschluss des Biologie-Studiums mit Diplom
Februar 1998 bis November 2002	Dissertation im Institut für Biochemie des Fachbereichs Biologie der Justus-Liebig-Universität Gießen unter Anleitung von Prof. Dr. A. Pingoud zum Thema „Entwicklung real-time-PCR-basierter Methoden für die moderne DNA-Analytik"
September 1998 bis August 2002	Stipendiat des Graduiertenkollegs „Molekulare Biologie und Pharmakologie" der Justus-Liebig-Universität Gießen

G. DANKSAGUNGEN

Vielen Dank

- Prof. Dr. Alfred M. Pingoud, für die Gewährung der Forschungsfreiräume, die ausgezeichneten Hilfestellungen, Unterstützungen und Förderungen, sowie die väterliche Führung mit fordernder aber liebevoller Hand.
- Prof. Dr. Manfred Kröger, für die exzellente Betreuung und das wissenschaftliche Rahmenprogramm des Graduiertenkollegs „Molekulare Biologie und Pharmakologie“. Mein Dank gilt auch der Deutschen Forschungsgemeinschaft für die Gewährung eines Doktorandenstipendiums im Rahmen des Graduiertenkollegs sowie den Kollegiaten.
- Dr. Meinhard Hahn, für Betreuung und Förderung, für die große Flexibilität und die gewährten Freiräume und das unbeschreiblich große persönliche Engagement, nicht nur auf wissenschaftlicher Ebene. Ich hoffe das Projekt „Menschwerdung 2002“ zu einem vertretbaren Abschluss gebracht zu haben…
- Dr. Annette Freist, Melanie Königshoff und „meinen“ Diplomanden und Praktikanten Anja Buch, Björn Tews, Christiane Thömmes, Hendrik Reuter, Patrick Schäfer und Steffen Kinzebach, für eure überaus engagierte und produktive Mitarbeit. Danke auch für den frischen Wind, den ihr in das Laborleben gebracht habt.
- unseren Kooperanden aus Wissenschaft und Wirtschaft: Dr. Annerose Anders, PD Dr. Burkhard Brandt, Dr. Jamilah Michel, Dr. Jürgen Sehrt, Dr. Johannes Hegemann für die DNA-Proben und Roland Kröhl, für die perfekte Kundenbetreuung und die unkomplizierte Hilfe bei Sonderwünschen, Drs. Karim Tabiti und Rudolf Seibl, für die finanzielle und technische Unterstützung die sie diesem Projekt zukommen ließen.
- Dr. Humaira Gowher, for your friendship, your helping hand, your inspiring discussions, and – last not least – your presence on the back seat of my motorbike. Labs seem to be empty without you.
- meinen lieben Kollegen im Institut, die ihr das Laborleben sehr bereichert habt. Mein besonderer Dank geht an Grischa Tödt, den coolsten aller Hacker, für seine Lehrbeispiele in Sozialisation (Hannover!) und die inspirierenden Gespräche, an Dr. Peter Friedhoff für die erlebenswerte Mitbetreuung von Teilen dieser Arbeit und an Ute Konradi, Ina Steindorf und Karina Urbach für ihre Hilfe und die erfrischende und liebenswerte Lebensart. Nicht zuletzt möchte ich auch Frau Backes danke sagen für so vieles: für die Naschereien, die Kürbissuppe, den gekochten Kaffee, die selbstgemachte Marmelade, …
- meinen Freunden, besonders Daniela Linder, Steffen Heinz, Klübers und Hußmanns, für den Mut, die Hoffnung, die Unterstützung und die Ablenkung die ihr mir gegeben habt.

Mein tiefster Dank gilt aber meinen Eltern, die mich bis an ihre Grenzen unterstützt und gefördert haben. Diese Arbeit ist nicht minder ihr Verdienst. Danke für alles!

H. ABKÜRZUNGSVERZEICHNIS

% (v/v)	Volumenprozent pro Volumen
% (w/v)	Gewichtsprozent pro Volumen
λ	Wellenlänge
μ	mikro (10^{-6})

A

A	Ampère
AAP	Agarosegel-Auftragspuffer
Abb.	Abbildung
APS	Ammoniumperoxodisulfat
ASCII	*American standard code for information interchange*
ATTC	*American Type Culture Collection*

B

Bp	Basenpaar(e)
BSA	bovines Serumalbumin
bzw.	beziehungsweise

C

c	centi (10^{-2})
°C	Grad Celsius
ca.	*circa*
CCD	*charge coupled device*
CE	Kapillarelektrophorese (*capillary electrophoresis*)
C_T	Schwellenwertzyklus (*threshold cycle*)
CV	Variationskoeffizient (*coefficient of variation*)

D

dATP	2'-Desoxyadenosin-5'-triphosphat
dCTP	2'-Desoxycytosin-5'-triphosphat
dGTP	2'-Desoxyguanosin-5'-triphosphat
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure (*deoxyribonucleic acid*)
dNTP	2'-Desoxyribonukleosid-5'-triphosphat
dsDNA	doppelsträngige DNA
dTTP	2'-Desoxythymidin-5'-triphosphat

E

eds.	Editoren (*editors*)
EDTA	Ethylendiamintetraacetat, Dinatriumsalz-Dihydrat
ELISA	*enzyme-linked immunosorbent assay*
EtBr	Ethidiumbromid
EtOH	Ethanol

F

FAM	5,6-Carboxyfluoreszein
FISH	Fluoreszenz-*in-situ*-Hybridisierung
FRET	Fluoreszenz-Resonanz-Energietransfer

G

g	Erdbeschleunigung oder Gramm
G	giga (109)
ggf.	gegebenenfalls
Gl.	Gleichung

H

h	Stunde(n)
HEX	5,6-Carboxy-4,7,2′,4′,5′,7′-hexachlorofluoreszein
HPLC	*high performance liquid chromatography*
HPSF	*highly purified salt-free*
Hrsg.	Herausgeber
Hz	Hertz (s-1)

K

k	kilo (10^3)
K	Kelvin oder Kopie(n)

L

l	Liter
LED	Licht emittierende Diode
LIF	Laser induzierte Fluoreszenz
LINE	*long intersperced nuclear elements*
LOH	*loss of heterozygosity*
LTR	*long terminal repeat*

M

m	Meter oder milli (10-3)
M	molar, Morgan oder Mega (106)
Mr	Molekolarmasse
MALDI	*matrix assisted laser desorption ionisation*
min	Minute(n)
MLE	*mariner-like elements*
mRNA	Boten-RNA (*messenger* RNA)
MW	Mittelwert

N

n	nano (10-9)
NaAc	Natriumacetat
NCBI	*National Center for Biotechnology Information*
Nt	Nukleotid(e)

O

OD optische Dichte
OOP Objekt-orientierte Programmierung

P

p piko (10^{-12})
PAGE Polyacrylamidgel-Elektrophorese
PCR Polymerase-Kettenreaktion (*polymerase chain reaction*)
POP *performance optimized polymer*

R

RFLP Restriktionsfragmentlängen-Polymorphismus
RFU relative Fluoreszenzeinheiten (*relative fluorescence units*)
RNA Ribonukleinsäure (*ribonucleic acid*)
rpm Umdrehungen pro Minute (*revolutions per minute*)
RT Raumtemperatur
RT-PCR Real-time-PCR

S

s Sekunde(n)
S. Seite
SD Standardabweichung(en) (*standard deviation*(s))
SDS Natriumdodecylsulfat (*sodium dodecyl sulphate*)
SINE *short interspersed nuclear elements*
SNP *single nucleotide polymorphism*
ssDNA einzelsträngige (*single stranded*) DNA

T

t Zeit (*time*)
T Temperatur
Tab. Tabelle
TAMRA Tetramethyl-6-carboxyrhodamin
Taq *Thermus aquaticus*
TE Tris-HCl-EDTA-Puffer
TEMED N,N,N',N'-Tetramethylethylendiamin
Temp. Temperatur
TET 5,6-Carboxy-4,7,2′,7′-tetrachlorofluoreszein
TIFF *Tagged image file format*
TLE tc1-like elements
Tm *Schmelzpunkt, Schmelztemperatur*
TPE *Tris-Phosphat-EDTA-Puffer*
Tris *Tris-(hydroxymethyl)-aminomethan*
Tris-HCl *Tris-(hydroxymethyl)-aminomethanhydrochlorid*

U

U Aktivitätseinheit für Enzyme (*unit*(s))
u. a. und andere oder unter anderem
UV Ultraviolett

V

v Volumen
V Volt
vgl. vergleiche

W

w Gewicht (*weight*)
W Watt
Wdh. Wiederhohlung(en)

Z

z. B. zum Beispiel

I. INHALTSVERZEICHNIS

1 EINLEITUNG

Man hat mir vorgeworfen, ich übertrage die Arbeitsweise der Physiker auf die Medizin. Nun gut, ich will mich gern schuldig bekennen. Denn heute liegt die Quelle des wissenschaftlichen Fortschritts mehr denn je an den Grenzen, in den Grauzonen zwischen den Spezialdisziplinen - und im Können und Mut jener Forscher, die sich dazwischen bewegen.

Manfred Baron von Ardenne

1.1 Die Entwicklung der molekularen DNA-Analytik

Die Erforschung der Krankheiten hat so große Fortschritte gemacht,
dass es immer schwerer wird, einen Menschen zu finden, der völlig gesund ist.

Aldous Leonard Huxley

Noch in den 1970er Jahren war die Desoxyribonukleinsäure (DNA) eines der am schwersten zu analysierenden Moleküle. Untersuchungen des Informationsgehalts der Basensequenz waren nur indirekt möglich, z. B. dem Informationsfluss folgend durch RNA- oder Protein-Sequenzierungen oder durch genetische Analysen an geeigneten Organismen. Mit der Entdeckung von Restriktionsendonukleasen wurde die direkte Untersuchung der DNA stark vereinfacht. Kartierungen von Restriktionsfragmentlängen erlaubten das Auffinden von Unterschieden im Erbgut verschiedener Individuen und eine Assoziation bestimmter Befunde auf der Ebene der DNA mit phänotypischen Ausprägungen wie z. B. erblichen Krankheiten.

Die Untersuchung der DNA vereinfachte sich erheblich mit dem Aufkommen einfacher technischer Verfahren zur Vervielfältigung definierter Abschnitte aus dem Erbmolekül, wodurch diese einer Analyse zugänglich wurden: Mit Hilfe der Polymerase-Kettenreaktion konnte innerhalb von wenigen Stunden aus nur einem einzigen DNA-Molekül praktisch jeder beliebige Abschnitt milliardenfach kopiert werden. Heute ist diese als „PCR" bezeichnete Methode wohl eines der wichtigsten, wenn nicht sogar das wichtigste biochemische Verfahren der molekularen DNA-Analytik und dessen unverzichtbare Grundlage.

In einer Vielzahl von Varianten wird die PCR nicht nur zur Vervielfältigung von Sequenzabschnitten verwendet, sondern auch zur Quantifizierung der Menge der in der Probe initial vorhandenen DNA-Abschnitte. Diese quantitativen Untersuchungen haben eine große Bedeutung bei Expressionsstudien, also der Analyse der Genaktivität (hierzu ist allerdings zunächst die reverse Transkription der Boten-RNA in DNA notwendig), sowie für die Tumordiagnostik, wo es während der Tumorprogression zum Verlust von Tumorsuppressorgenen oder der Vervielfältigung der Kopienzahl von (Proto-)Onkogenen kommen kann. Nicht zuletzt erlaubt die quantitative DNA-Analytik auch die Bestimmung von Keimbelastungen oder Virustitern.

Mit der Sequenzierung des humanen Genoms ist deutlich geworden, dass sich individuelle Genome nicht in wenigen, größeren Abschnitten unterscheiden, sondern in vielen, dafür aber sehr kleinen, meist nur ein Nukleotid umfassenden Abschnitten. Diese erblichen Unterschiede tragen zur phänotypischen Variation bei, beeinflussen also die individuelle anthropometrische Charakteristik, das Risiko für Erkrankungen sowie die Reaktion auf Umwelteinflüsse. Die einfache Typisierung dieser polymorphen Loci bei vielen Individuen ermöglicht zunächst die Assoziation bestimmter Genotypen mit individuellen Erkrankungsrisiken oder Medikamenten(un-)verträglichkeiten und später deren einfache und schnelle Diagnose, birgt aber gleichzeitig auch das Risiko einer wertenden Klassifizierung untersuchter Individuen.

Die vorliegende Arbeit beschreibt die Entwicklung neuer PCR-basierter Verfahren zur einfachen, schnellen und hoch-präzisen Quantifizierung und Typisierung genetischer Elemente sowie neuer Algorithmen und deren Einsatz in der automatisierten Auswertung der Messdaten mit Hilfe computergestützter Datenverarbeitung.

1.2 Die Polymerase-Kettenreaktion

The greatest weakness of men is our impossibility of a real understanding of the exponential function.

Albert Allen Bartlett

Mit Hilfe der Polymerase-Kettenreaktion (PCR) ist es möglich, spezifische Sequenzbereiche aus geringsten Ausgangsmengen von Nukleinsäuren *in vitro* und zudem schnell zu amplifizieren, um sie so einer Analyse oder Weiterverarbeitung zugänglich zu machen. Das Grundprinzip der Polymerase-Kettenreaktion ist in **Abb. 1-2** dargestellt. Ein doppelsträngiges DNA-Molekül wird durch Hitzeeinwirkung aufgeschmolzen (denaturiert). Die Einzelstränge dienen in der Folge als Matrize für die enzymatisch katalysierte Polymerisation von Desoxyribonukleotiden, wodurch wieder doppelsträngige DNA-Moleküle entstehen. Die als Primer bezeichneten Oligodesoxyribonukleotide definieren dabei den zu kopierenden Sequenzabschnitt, indem sie an Orten komplementärer Sequenz mit der Ziel-DNA hybridisieren und als Starter für die Polymerisation dienen. Der Prozess exponentieller Produktbildung wird von verschiedenen Faktoren begrenzt. Im Laufe der PCR geht die Netto-Produktbildung daher schließlich auf Null zurück und die Gesamtmenge an PCR-Produkt erreicht einen Plateauwert. Die Ursachen dafür werden z. B. bei MORRISON & GANNON umfassend diskutiert.[1]

Erstmals beschrieben wurde ein Verfahren zur biochemischen Amplifikation von DNA *in vitro* bereits 1971 von KLEPPE *et al*.[2] Die damals verfügbare KLENOW-Polymerase zur DNA-Synthese war nicht hitzestabil und musste nach jeder Denaturierung erneut dem Reaktionsansatz zugegeben werden. Dieser Aufwand stand einer weiten Ver-reitung dieser Methode entgegen, bis MULLIS 1988 den genialen und durch seine Einfachheit bestechenden Einfall hatte, eine aus thermophilen Bakterien (*Thermus aquaticus*) isolierte DNA-Polymerase zu verwenden, welche den Denaturierungsprozess der DNA unbeschadet übersteht.[3]

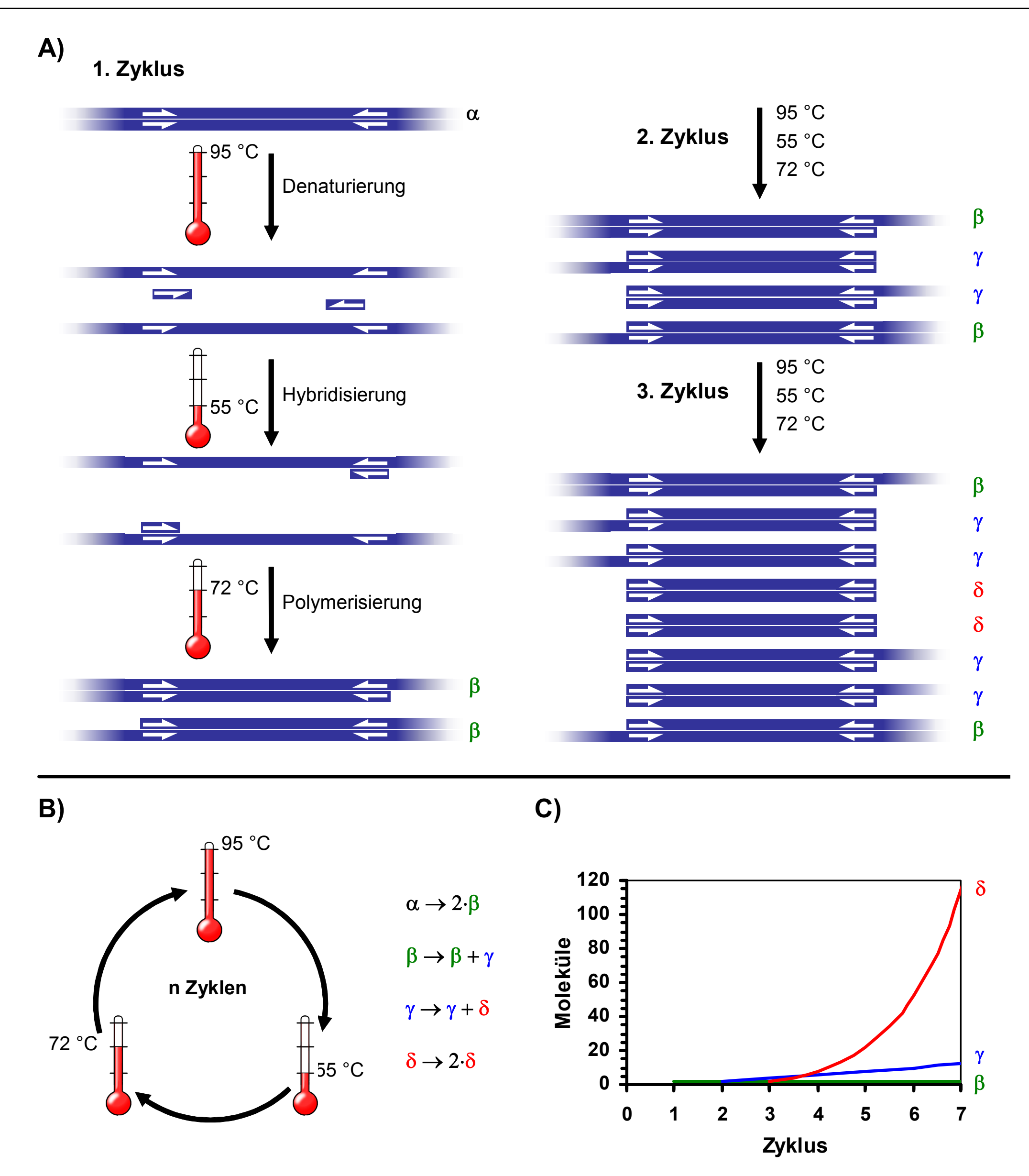


Abb. 1-1 **Schematische Darstellung der PCR. A)** Die dsDNA α mit der Zielsequenz wird durch Hitzeeinwirkung in die Einzelstränge aufgeschmolzen. Durch Absenken der Temperatur hybridisieren die Primer bzw. mit den zu ihnen komplementären Abschnitten. Während der anschließenden, durch die thermostabile DNA-Polymerase katalysierten Polymerisationsreaktion bei einer Temperatur zwischen 65°-72° C werden die Primer verlängert, und es entstehen wieder doppelsträngige Produkte. **B)** Durch Wiederholungen des Temperaturprofils nimmt die Gesamt-Produktmenge exponentiell zu. Aus α entstehen dabei zwei Produkte β, die jeweils nur einseitig in einem Strang definiert terminiert sind. Die Produkte β bilden ihrerseits zusätzlich ein Produkt γ, welches einseitig in beiden Strängen definiert terminiert ist. **C)** Hieraus entsteht in den weiteren Zyklen das gewünschte, beidseitig terminierte Produkt δ, welches schließlich durch Verdoppelung in den nachfolgenden Zyklen exponentiell in seiner Menge zunimmt.

Damit wurde eine schnelle, einfache und automatisierbare *in-vitro*-Amplifikation definierter DNA-Fragmente aus geringen Mengen heterogener DNA möglich. Diese technische Verbesserung hatte überaus großen Einfluss auf die gesamte weitere Entwicklung der Molekularbiologie und legte den Grundstein für molekulardiagnostische Verfahren in der Medizin. Das Verfahren wurde von der Firma Cetus patentiert, die ihre Rechte für jedwede kommerzielle Anwendungen an das Schweizer Pharmaunternehmen Hoffman-La Roche verkaufte.

Seither wurde eine nahezu unüberschaubare Vielzahl von PCR-Varianten entwickelt. Heute ist die PCR eine der wichtigsten Methoden in der molekularen Biologie und molekularen Medizin. Ihre ersten Anwendungen waren der Mutationsnachweis im Genom von Patienten mit Beta-Thalassämie [4] und Sichelzellanämie [5] sowie der Nachweis proviraler DNA in Leukozyten.[6] Heute findet sie Verwendung in einem überaus breiten thematischen Spektrum, z. B. beim Nachweis von Viren oder Keimen, bei der Sequenzierung, dem Verwandtschaftsnachweis, der Erstellung von Transkriptionsprofilen und der Quantifizierung von Nukleinsäuren.[7-10] Zudem lassen sich mit Hilfe der PCR in einfacher Weise beliebige Sequenzabschnitte des Nukleinsäurebestandes eines Organismus klonieren. Die Vielzahl entwickelter PCR-Varianten ermöglicht u. a. eine zielgerichtete oder zufällige Veränderung der DNA-Sequenz sowie sogar die Synthese größerer, in dieser Form zuvor nicht existenter Sequenzabfolgen. Eine Übersicht zu Anwendungen der PCR-Technik geben spezielle Übersichtsartikel,[7-11] Bücher [12-14] sowie die Zeitschrift *PCR-Methods and Applications* (nicht mehr aufgelegt).

1.2.1 Quantitative PCR

Eine bedeutende Anwendung der PCR ist die Quantifizierung von bestimmten Nukleinsäuren in einer Probe. Die in den Proben vorhandenen Mengen an Nukleinsäuren sind in der Regel so gering (von einem Molekül bis zu einigen Millionen Molekülen), dass ihre direkte Quantifizierung nur mit unverhältnismäßig großem Aufwand oder auch gar nicht möglich wäre. Die PCR bietet eine einfache Möglichkeit, die Menge der zu quantifizierenden Sequenzen durch ihre Amplifikation einer quantitativen Analyse zugänglich zu machen, da sich die Akkumulation der Produkte anfänglich durch einfache mathematische Beziehungen beschreiben lässt.

Bei der PCR handelt es sich um einen Prozess zyklisch wiederholter Produktakkumulation. In jedem Amplifikationszyklus wird die Menge an PCR-Produkt um einen Faktor E_i ($0 \leq E_i \leq 1$) erhöht (der Index bezeichnet die Zyklus-Nummer). Daraus ergibt sich die jeweils folgende Produktmenge N durch

$$N_{i+1} = \left(1 + E_i\right) \cdot N_i \qquad \text{Gl. 1.1}$$

und die Gesamt-Produktmenge nach n Zyklen und N_0 initialen Kopien der Zielsequenz durch

$$N_n = N_0 + \sum_{i=1}^{n} N_{i-1} \cdot E_i \qquad \text{Gl. 1.2}$$

Diese Funktion kann auch als iteratives Produkt geschrieben werden:

$$N_n = N_0 \cdot \prod_{i=1}^{n} (1 + E_i) \qquad \text{Gl. 1.3}$$

wobei der Ausdruck $(1+E_i)$ auch als Effizienz ε_i $(1 \leq \varepsilon_i \leq 2)$ bezeichnet wird. Für konstante Effizienzen ($\varepsilon_i = \varepsilon =$ const.) entspricht **Gl. 1.3** dem exponentiellen Wachstumsgesetz

$$N_n = N_0 \cdot \varepsilon^n \qquad \text{Gl. 1.4}$$

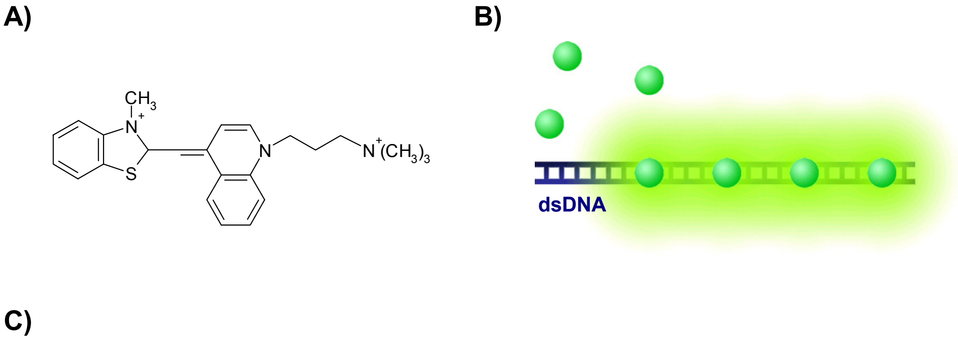
Diese Vereinfachung gilt nur in der frühen, so genannten exponentiellen oder log-linearen Phase der Reaktion, wenn limitierende Einflüsse noch vernachlässigbar sind. Später bewirken die Limitationen eine fortschreitende Abnahme der Amplifikationseffizienz, die in der späten Plateau-Phase der PCR schließlich auf einen Wert von eins abfällt.

Nach einer hinreichenden Anzahl Zyklen kann die Produktmenge mit verschiedenen Methoden (z. B. UV-spektroskopisch oder densitometrisch) bestimmt werden. Da die Amplifikationseffizienzen ε_i nicht *a priori* bekannt sind, kann N_0 allerdings nicht aus einer alleinigen Bestimmung von N_n berechnet werden. Zur Lösung dieses Problems stehen im Wesentlichen drei unterschiedliche Methoden zur Verfügung:

- **Limitierende Verdünnungstechnik.** Hierbei werden serielle Verdünnungen der Probe in die PCR eingesetzt und bestimmt, bei welcher Verdünnungsstufe sich im Mittel nur noch eine einzige Kopie der Zielsequenz im Reaktionsansatz befindet. Das ist dann der Fall, wenn in der Hälfte der Reaktionsansätze der betreffenden Verdünnung ein spezifisches PCR-Produkt nachgewiesen werden kann. Die Anzahl der benötigten Reaktionsansätze steigt schnell mit den Anforderungen an die Genauigkeit der Bestimmung. Ein großes Problem stellen die in Reaktionsansätzen mit sehr geringen initialen Templatkopienzahlen bevorzugt entstehenden unspezifischen PCR-Produkte („Primerdimere") dar, weiterhin machen selbst geringste Kontaminationen mit Templaten (z. B. PCR-Produkte) in den Reaktionsansätzen die Ergebnisse unbrauchbar.[15,16]

- **Kompetitive PCR.** Sehr verbreiteten Einsatz fanden Koamplifikationstechniken, bei welchen dem Reaktionsansatz ein interner Standard zugesetzt wird, der ko-amplifiziert wird. Unter der Voraussetzung, dass der Kompetitor mit der identischen Effizienz amplifiziert wird wie das Probentemplat, sind damit sehr präzise Quantifizierungen möglich. Dazu müssen Kompetitor und Probe möglichst ähnliche Sequenzen aufweisen, aber dennoch analytisch unterscheidbar sein. Ist diese Voraussetzung erfüllt, so bleibt das Verhältnis der Produkte von Probe und Kompetitor während der gesamten Reaktion, also auch in der Plateauphase, konstant. Anhand des Mengenverhältnisses der Produkte nach der PCR und der Kenntnis über die eingesetzte Menge Kompetitor lässt sich so die initiale Menge der Probe bestimmen. Mögliche Nichtlinearitäten der Methoden zur Produktquantifizierung (z. B. Densitometrie) können wirksam ausgeglichen werden, indem in einer Titrationsreihe diejenige initiale Menge Kompetitor bestimmt wird, bei welcher die Produktmengen von Probe und Kompetitor gleich sind (Äquivalenzpunkt). Nachteile der quantitativen kompetitiven PCR sind der relativ hohe Verbrauch an Probenmaterial sowie die meist zeitaufwändige Post-PCR-Analytik.[17-21]

Real-time-PCR. Ein weiteres Verfahren macht sich den Umstand zunutze, dass die Amplifikationseffizienzen in der exponentiellen Phase der PCR näherungsweise konstant sind. Da hierzu bestimmt werden muss, bis zu welchem Amplifikationszyklus die Produktzunahme noch exponentiell ist, muss diese bei jeder individuellen Reaktion verfolgt werden. Eleganterweise geschieht dies störungsfrei über die Messung von Fluoreszenzlicht, dessen Intensität proportional zur Produktmenge ist. Da diese Messung in Echtzeit während der PCR erfolgt, wird dieses Verfahren als Real-time-PCR bezeichnet.[22-24]



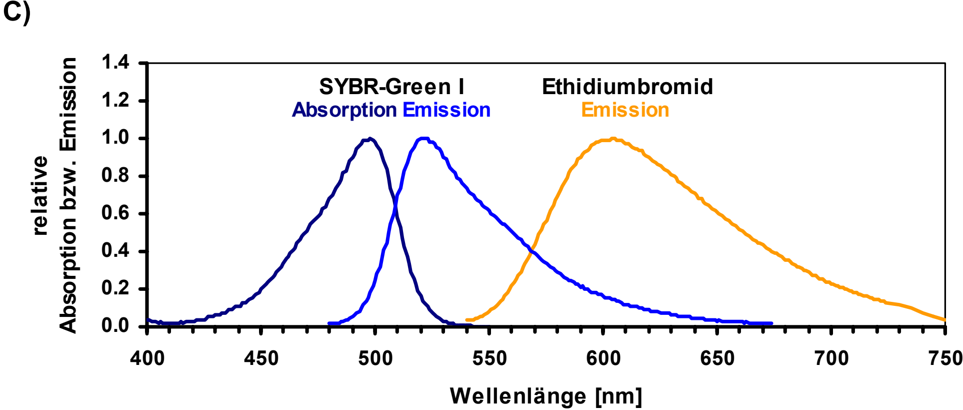


Abb. 1-2 Struktur und Eigenschaften von SYBR Green I. A) Wahrscheinliche Struktur von SYBR Green I. Entnommen aus der Patentschrift #5321130 der US-Patentbehörde, Table 1, Compound 4. **B)** Generierung von Fluoreszenzsignal während der PCR durch SYBR-Green I. Das Anregungslicht wird von frei in Lösung befindlichem SYBR-Green I (●) nicht absorbiert. Nur der in der kleinen Furche der dsDNA gebundene Farbstoff fluoresziert nach Anregung. **C)** Absorptions-/Emissionsspektren von SYBR Green I. Die Spektren einer 1:30000-fachen Verdünnung von SYBR Green I in einer Lösungen mit 10 ng/μl genomischer DNA in Tris/HCl, pH 8.5 bei 20°C, wurden aufgezeichnet wie in 2.7 beschrieben. Zum Vergleich ist das Emissionsspektrum von Ethidiumbromid nach Anregung mit Licht einer Wellenlänge von 320 nm unter sonst gleichen Bedingungen gezeigt. Die Konzentration von Ethidiumbromid in der gemessenen Lösung betrug 10 mg/l.

1.2.2 Real-time-PCR

Bei der Real-time-PCR (RT-PCR) wird die Produktakkumulation in Echtzeit verfolgt, um die exponentielle Phase der Reaktion zu identifizieren und in der Folge für eine Quantifizierung der Templatmenge zu nutzen. Die Messung der Produktzunahme geschieht störungsfrei anhand von Fluoreszenzlicht, welches in Abhängigkeit der Produktmenge generiert wird.[22,23]

Eine Möglichkeit zur Generierung produktabhängiger Fluoreszenz bieten dsDNA-bindende Fluoreszenzfarbstoffe. Der hierfür zuerst verwendete Farbstoff ist Ethidiumbromid,[25-27] der von UV-Licht (254–366 nm) angeregt werden kann und ein Emissionsmaximum bei 590 nm besitzt (Emissionsspektrum siehe **Abb. 2-3**). Neben Ethidiumbromid wurden auch andere fluoreszierende Interkalatoren verwendet, wie z. B. YO-PRO-1.[28,29] Der heute für RT-PCRs üblicherweise verwendete dsDNA-sensitive Farbstoff ist SYBR Green I:

SYBR-Green I ist ein unsymmetrischer Cyaninfarbstoff, strukturell eng verwandt mit den dimeren DNA-sensitiven Fluoreszenzfarbstoffen YOYO-1 und TOTO-1.[30,31] Die genaue Struktur von SYBR Green I wird von den Lizenzinhabern geheim gehalten. Aus den Patentschriften #05321130 und #5658751 der US-Patentbehörde (siehe http://www.uspto.gov) geht entsprechend der Emissionseigenschaften jedoch hervor, dass der Farbstoff wahrscheinlich die in Abb. 1-2 A gezeigte Struktur aufweist. Der Farbstoff interkaliert nicht, sondern bindet weitgehend sequenzunspezifisch in der kleinen Furche doppelsträngiger DNA, mit einer etwa 100mal höheren Affinität als Ethidiumbromid. Das in der kleinen Furche gebundene SYBR-Green I fluoresziert nach Anregung etwa 1000mal stärker als der freie Farbstoff, weshalb SYBR Green I sehr gut geeignet ist, die Akkumulation doppelsträngige PCR-Produkte in der RT-PCR sichtbar zu machen (Abb. 1-2 B).[26,27] SYBR Green I lässt sich, in Gegenwart von doppelsträngiger DNA, mit Blaulicht (480 nm) anregen und zeigt ein dem Fluoreszein vergleichbares Emissionsspektrum mit einem Maximum bei 520 nm bei einer Quantenausbeute von etwa 0.8 (Abb. 1-2 C).

Durch die Verwendung fluorophormarkierter Oligonukleotide als Sonden, die auf der Zielsequenz hybridisieren, lassen sich PCR-Produkte auch sequenzspezifisch detektieren.[32-34] Hierbei wird entweder die produktabhängige Abnahme der Löschung (Quench) einer Reporter-Fluoreszenz oder die Zunahme des Resonanzenergietransfers von einem Donor- auf einen Akzeptorfluorophor gemessen.[35-37] Der Fluoreszenz-Resonanzenergietransfer (FRET) ist ein spektroskopischer Prozess, bei dem Energie strahlungsfrei durch Dipol-Dipol-Welchselwirkungen zwischen Molekülen mit überlappenden Emissions- und Absorptionsspektren über einen Bereich von 10-100 Å übertragen wird.[36] Zu den bekanntesten Varianten von Real-time-Detektionssystemen, die FRET zur Generierung sequenzspezifischer Fluoreszenzsignale nutzen, zählen:

● **Hybridisierungssonden** (**Abb. 1-3 A**). Das Sondensystem besteht hier aus einem Paar terminal markierter Sonden, die an unmittelbar benachbarten Stellen auf der Zielsequenz hybridisieren, so dass der Donorfluorophor der einen Sonde in räumlicher Nähe des Akzeptorfluorophors der anderen Sonde positioniert ist. Zur Anregung wird Licht einer Wel-

lenlänge verwendet, welches nur vom Donorfluorophor, nicht jedoch vom Akzeptorfluorophor absorbiert wird. Befinden sich Donor- und Akzeptorfluorophor in direkter Nachbarschaft, ist ein Fluoreszenz-Resonanzenergietransfer vom angeregten Donor- auf den Akzeptorfluorophor möglich, welcher die übertragene Energie schließlich als Fluoreszenzlicht abgibt.[26]

Hydrolysesonden (**Abb. 1-3 B**). Die auch als TaqMan-Sonden bezeichneten Oligonukleotide sind 5′-terminal mit einem Reporterfluorophor und intern mit einem Quencher markiert. Während der Verlängerung eines Primers trifft die hier verwendete *Taq* DNA-Polymerase auf die Sonde, die durch die endonukleolytische Aktivität des Enzyms hydrolysiert wird. Dabei wird der Reporterfluorophor räumlich vom Quencher getrennt und fluoresziert bei Anregung.[23,38,39]

Molecular Beacons (**Abb. 1-3 C**). Diese Sonden sind beidseitig terminal modifiziert. Sie tragen an dem einen Ende einen Reporterfluorophor und an dem anderen Ende einen Quencher. Ihre terminalen Sequenzen sind zueinander komplementär, so dass diese Sonden in freier Lösung in einer charakteristischen *stem-loop*-Struktur vorliegen. Die *loop*-Region ist spezifisch für die Zielsequenz. Sie hybridisiert während der Primerbindungsphase mit dem komplementären Strang des PCR-Produkts und öffnet dabei die *stem-loop*-Struktur, wodurch Quencher und Reporterfluorophor räumlich getrennt werden. Als Folge hiervon kann der Reporterfluorophor nach Anregung fluoreszieren.[40-42]

Sunrise Primer (**Abb. 1-3 D**). Bei diesen Sonden handelt es sich um Primer mit selbstkomplementären Sequenzabschnitten am 5′-Terminus, der in Lösung eine *hairpin*-Struktur aufweist. Die kommerziell als Amplifluor Hairpin Primer bezeichneten Sonden sind, vergleichbar den Molecular Beacons, mit Quencher und Reporterfluorophor markiert, die durch die *hairpin*-Struktur in enger räumlicher Nachbarschaft zueinander gehalten werden. Der 3′-Terminus ist komplementär zur Zielsequenz und dient hier als Primer. Er wird nach der Primerbindungsphase verlängert, womit die Sonde in das PCR-Produkt eingebaut ist. Im darauffolgenden Amplifikationszyklus dient dieses Produkt als Matrize für die Verlängerung des Gegenprimers. Im gebildeten doppelsträngigen PCR-Produkt liegt die *hairpin*-Struktur geöffnet vor, Quencher und Reporterfluorophor sind also räumlich getrennt, weshalb der Reporterfluorophor nach Anregung fluoresziert. Problematisch bei diesem System ist die Generierung von Fluoreszenzsignalen durch die Bildung von Primerdimeren.[24,43,44]

Scorpion Primer (**Abb. 1-4**). Dieses System ist eine Variante der Sunrise Primer, welches die Bildung fluoreszierender Primerdimere unterbindet. Erreicht wird dies durch den Block der Verlängerung des Gegenstranges in den Primer hinein. Das 5′-terminale Ende dieses Primers, welches in der Ausbildung der *stem-loop*-Struktur der frei vorliegenden Form beteiligt ist, ist komplementär zu einem Sequenzabschnitt des durch die Verlängerung diesen Primers entstandenen Einzelstranges und kann unter Auflösung der *stem-loop*-Struktur auf diesem Strang binden.

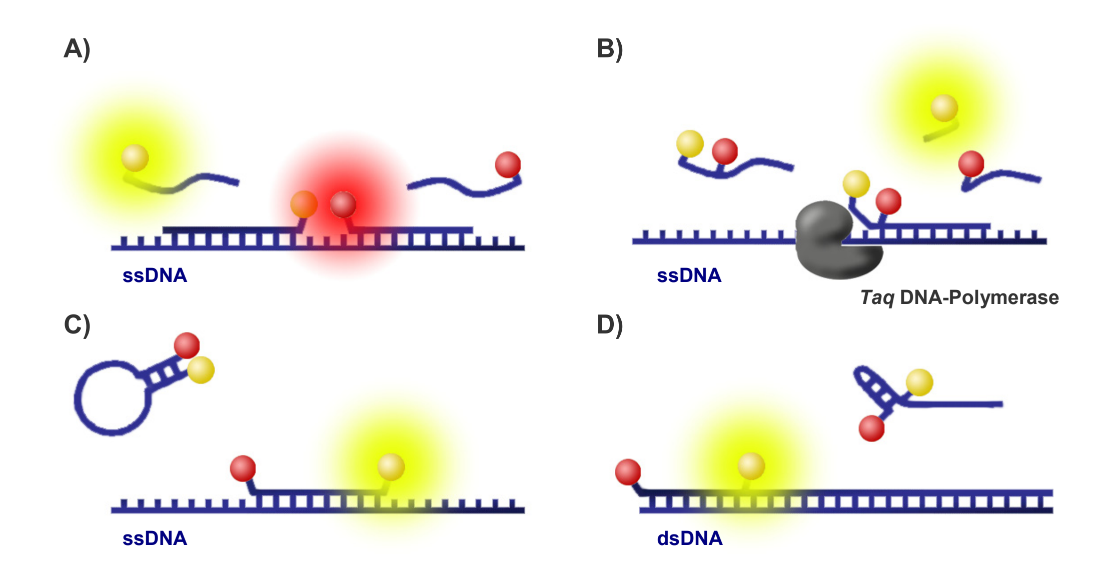


Abb. 1-3 Prinzipien zur Generierung sequenzspezifischer Fluoreszenzsignale während der PCR. Gezeigt sind schematische Darstellungen einiger ausgewählter Verfahren zur Generierung amplikonspezifischer Fluoreszenzsignale, die in der RT-PCR zur Anwendung kommen. **A)** Hybridisierungssonden. **B)** Hydrolysesonden. **C)** Molecular Beacons. **D)** Sunrise Primer. (gelb) Donor- bzw. Reporterfluorophor, (rot) Akzeptorfluorophor bzw. Quencher. Einzelheiten zu den Prinzipien der Signalgenerierung finden sich im Text.

Neben Oligonukleotiden finden auch die strukturell ähnlichen *peptide nucleic acids* (PNAs) Verwendung in der sequenzspezifischen Detektion: Die sogenannten *light-up* Sonden sind PNAs mit Thiazolorange als Fluorophor. Bei der Hybridisierung mit Nukleinsäuren unter Ausbildung von Duplex- oder Triplexstrukturen erhöht sich die Fluoreszenzintensität dieser Sonden beträchtlich, was zur Detektion genutzt wird. Allerdings steigt die unspezifische Fluoreszenz im Laufe der PCR, weshalb die Sensitivität geringer ist als die der oben genannten FRET-basierten Sondensysteme.[45-47]

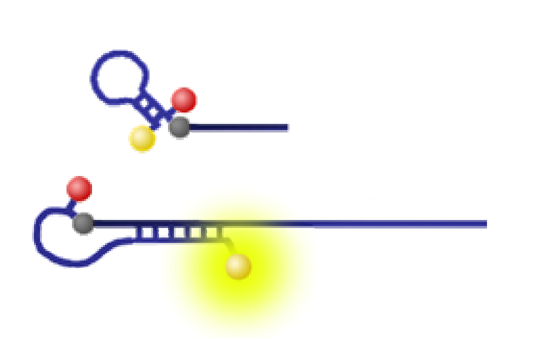


Abb. 1-4 Funktionsweise der Scorpion Primer. Der freie Primer bildet eine *Stem-loop*-Struktur, in welcher die Fluoreszenz des Reporters (gelb) durch einen Quencher (rot) gelöscht wird. Das Öffnen der *Stem-loop*-Struktur erfolgt nach Bindung des 5′-terminalen Abschnitts des Primers auf dem DNA-Strang, der durch die Verlängerung seines 3′-Teminus entstanden ist. Eine Öffnung des Primers infolge der Bildung von Primerdimeren oder unspezifischen Produkten wird verhindert durch den Block der Verlängerung der Gegenstranges durch eine interne Hexethylenglykolgruppe (grau).

Einige Verfahren messen statt der Zunahme eine Abnahme der Fluoreszenz bzw. eine Zunahme des Quenchs. Dabei wird die Fluoreszenz der hybridisierten Sonde durch einen Guaninrest im PCR-Produkt gelöscht.[48,49] Die Sonden sind einfach zu konstruieren und zu synthetisieren und stellen damit eine preisgünstige Möglichkeit der sequenzspezifischen Detektion von PCR-Produkten dar, allerdings ist die sehr geringe Abnahme der Fluoreszenz in den frühen, exponentiellen Zyklen der PCR, wo die freien Sonden noch im großen Überschuss vorliegen, nur schwer zu messen.

Die absolute Quantifizierung der initialen Templatkopienzahlen unbekannter Proben erfolgt bei der RT-PCR anhand externer Standards mit verschiedenen bekannten initialen Kopienzahlen des Templats, die idealerweise unter identischen Bedingungen wie die Proben amplifiziert werden. Da in allen Fällen die Phase exponentieller Signalzunahme identifiziert werden kann, lässt sich ein Signal-Schwellenwert angeben, bei welchem sich jede Einzelreaktion in der exponentiellen Phase befindet. Der Zeitpunkt (in Zyklen), zu dem das Signal diesen Schwellenwert erreicht, wird als Schwellenwertzyklus, *crossing point* oder C_T-Wert bezeichnet. Der genaue Wert wird durch Interpolation der Messdaten gewonnen und muss keine Ganzzahl sein. Unter der Voraussetzung, dass die Signalintensität S proportional zur Produktmenge ist (mit p als Proportionalitätsfaktor), folgt aus Gleichung **Gl. 1.4** für die Produktmenge am Schwellenwertzyklus

$$S = p \cdot N_{C_T} = p \cdot N_0 \cdot \varepsilon^{C_T} \qquad \text{Gl. 1.5}$$

Die nach C_T aufgelöste Form

$$C_T = -\frac{1}{\log(\varepsilon)} \cdot \log(N_0) - \frac{\log(p) - \log(S)}{\log(\varepsilon)} \qquad \text{Gl. 1.6}$$

ist die Gleichung der Kalibrierfunktion. Der einfache lineare Zusammenhang zwischen dem C_T-Wert und dem Logarithmus der initialen Templatmenge lässt sich erkennen, wenn $m = -(\log(\varepsilon))^{-1}$ und $b = m \cdot (\log(p) - \log(S))$ gesetzt wird:

$$C_T = m \cdot \log(N_0) + b \qquad \text{Gl. 1.7}$$

Wenn S, p und ε für alle Standards (und Proben) gleich sind, können die Parameter m und b der Kalibriergeraden durch eine Regressionsanalyse der C_T-Werte der Standards gegen die Logarithmen ihrer bekannten Konzentrationen ermittelt werden. Die initiale Templatmenge N_0 einer Probe lässt sich dann anhand ihres C_T-Wertes aus **Gl. 1.7** berechnen:

$$N_0 = 10^{\frac{C_T - b}{m}} \qquad \text{Gl. 1.8}$$

Der Vorteil dieses Verfahrens liegt in der Robustheit der Bestimmung der Effizienz. Diese wird hier indirekt über die C_T-Wert-Unterschiede der Standards bestimmt, womit die Produktakkumulation seit dem ersten Amplifikationszyklus berücksichtigt wird. Dies ist bei der Bestimmung der Effizienz über die messbare exponentielle Phase einer einzelnen Probe alleine nicht möglich; hier können nur die Punkte der Zyklen verwendet werden, bei denen schon eine messbare Produktakkumulation stattgefunden hat, der Übergang in die Plateauphase aber noch nicht erreicht ist. Das sind bei der heutigen Messtechnik ca. 3-6

Punkte (statt 20-40 Punkten bei der oben beschriebenen Methode), die zur Bestimmung der Effizienz herangezogen werden können, was einen vergleichsweise großen Fehler mit sich bringt. Die Präzision der Bestimmung von N_0 bei der RT-PCR ist sehr von der Präzision der Bestimmung der C_T-Werte abhängig, da diese fehlerbehaftet ist und im Exponenten steht (**Gl. 1.8**). So führt bei einer Amplifikationseffizienz von 2 ein Fehler von nur 0.14 Zyklen bei der Bestimmung des C_T-Werts einer Probe schon zu einem Fehler von mehr als 10 % bei der Bestimmung der initialen Templatmenge. Die exakte Bestimmung der C_T-Werte ist daher entscheidend für die Quantifizierung mit RT-PCR-Verfahren.

1.3 Genotypisierung und Mutationsdetektion

These individual differences are highly important for us, as they afford materials for natural selection to accumulate [...].

Charles Darwin in The Origin of Species

Genotypisierung bezeichnet auf molekulargenetischer Ebene die Identifizierung der Sequenz definierter, kurzer Abschnitte der Erbinformation, die interindividuelle Variationen aufweisen können. Die unterschiedlichen Varianten entstehen durch Mutationsereignisse, weshalb Mutationsdetektion und Genotypisierung methodisch prinzipiell dasselbe sind. Etablieren sich Mutationen in der Keimbahn, kann sich die Sequenzvariante im Genpool der Population anreichern – vorausgesetzt, es wirkt kein stark negativer Selektionsdruck auf den korrespondierenden Phänotyp, der erblich bedingten Merkmalsausprägung. Die Varianten dieser vererbten, polymorphen Sequenzabschnitte werden auch als Allele bezeichnet, weshalb statt Genotypisierung auch das Synonym Allelotypisierung verwendet wird.

Das DNA-analytische Verfahren der Genotypisierung erlaubt somit die Identifizierung von Arten, Populationen, Individuen, Genen bzw. Allelen sowie von Änderungen der genetischen Information. Die Anwendungsgebiete in der Genotypisierung sind entsprechend breit gefächert und umfassen die Diagnose von erblich bedingten Krankheiten,[50] erblichen Prädispositionen,[51] forensische Analysen,[52] Elternschaftsnachwiese,[53] Abstammungsanalysen,[54] die Identifizierung von Mikroorganismen, insbesondere von Krankheitserregern,[18,55] bis hin zur Aufklärung evolutionärer Zusammenhänge.[56]

Bei den Polymorphismen kann es sich u. a. um Insertionen oder Deletionen von Sequenzabschnitten oder einzelnen Basenpaaren handeln. Diese können in der Regel sehr einfach nachgewiesen werden, indem der polymorphe Sequenzbereich mittels PCR amplifiziert wird und anschließend die Fragmentlängen der PCR-Produkte analysiert werden. Eine analytisch wichtige Gruppe von Längen-Polymorphismen bilden die *short tandem repeats* (STRs) mit tandemartig wiederholten Sequenzabschnitten einer Grundlänge von ein bis fünf Basenpaaren.[54,57] Die repetitive Anordnung der Grundeinheiten begünstigt Mutationsereignisse, die zu einer Änderung der Anzahl von Wiederholungen führen. Der genetische Fingerabdruck zur Identifizierung von Individuen beruht heute hauptsächlich auf der STR-Analyse.[52,53]

Deutlich schwieriger nachzuweisen sind Sequenz-Polymorphismen, die in den meisten Fällen als Folge einer Punktmutation nur ein einzelnes Basenpaar umfassen. Die Analyse solcher Einzelnukleotid-Polymorphismen (*single nucleotide polymorphisms*, SNPs) hat aktuell sehr an Bedeutung gewonnen, da ihre genomische Dichte von keiner anderen Klasse genetischer Marker erreicht wird und die meisten SNPs in nicht-kodierenden Sequenzen liegen und daher wahrscheinlich phänotypisch neutral sind.[58,59] Dennoch sind sehr viele SNPs mit komplexen Erkrankungen assoziiert oder modulieren Veranlagungen für Krankheitsanfälligkeiten oder die Verträglichkeit bestimmter Medikamente.[60] Die Zahl bekannter humaner SNPs nimmt immer noch beständig zu. Die International SNP Map Working Group hatte im Jahre 2001 bereits 1.4 Millionen SNPs im humanen Genom kartiert.[59] Schätzungen gehen davon aus, dass sich die Genome zweier Individuen im Mittel in ca. 3.2 Millionen Positionen unterscheiden (~0.5 ‰ der Erbinformation).[61]

Seit den ersten Entwicklungen zur nukleinsäurebasierten Genotypisierung in den 1970er Jahren mit der Entdeckung der Restriktionsfragmentlängen-Polymorphismen (RFLPs)[62] gewann die Genotypisierung durch die Entwicklung der PCR rasant an Bedeutung. Ihre Anwendung intensivierte sich anschließend noch durch die Einbindung in Hochdurchsatz-Verfahren, wie sie insbesondere im Humanen Genomprojekt eingesetzt wurden. Heute sind weit mehr als 100 Verfahren zur Genotypisierung beschrieben, wobei sich die den Methoden zugrunde liegenden Prinzipien zur Detektion bzw. Identifizierung von Sequenzvariationen allerdings in wenige Gruppen einteilen lassen. Die unmittelbare Identifizierung der polymorphen Sequenz wird durch Techniken erreicht, die einen Sequenzierungsschritt umfassen (*sequencing, minisequencing, template-directed dye-terminator incorporation*).[63-67]

Heteroduplex-Analysen machen sich den Umstand zunutze, dass dsDNA-Fragmente mit Basenfehlpaarungen eine veränderte Mobilität in elektrophoretischen oder chromatographischen Trennverfahren aufweisen.[68] Die Sensitivität der Trennungen kann durch besondere Gradiententechniken wie die denaturierende Gradienten-Gelelektrophorese (DGGE),[69,70] die Temperaturgradienten-Gelelektrophorese (TGGE)[71,72] oder die temperaturmodulierte Hochleistungs-Flüssigkeitschromatographie (TmHPLC)[73-75] verbessert werden.

Andere Verfahren nutzen die enzymatische Erkennung von Basenfehlpaarungen in Heteroduplices zum Nachweis von Sequenzvarianten. Die Erkennung erfolgt dabei z. B. durch die T4 Endonuklease VII,[76,77] die T7 Endonuklease I[78] oder DNA-Reparatur-Enzyme wie z. B. MutS.[79,80] Liegt die polymorphe Position in der Erkennungssequenz einer beliebigen Restriktionsendonuklease, so können Sequenzvarianten durch Spaltung mittels solcher sequenzspezifischer Nukleasen nachgewiesen werden.[81-83]

Ligationsbasierte Methoden wie das *oligonucleotide ligation assay* (OLA),[84-87] die Ligase-Kettenreaktion (*ligase chain reaction*, LCR)[88-90] oder die *dye-labeled oligonucleotide ligation* (DOL)[91] beruhen auf dem hohen Diskriminierungspotenzial von Ligasen, die eine Bruchstelle eines Stranges in der dsDNA („*nick*") praktisch nicht ligieren, wenn an der Ligationsstelle eine Basenfehlpaarung vorliegt.

Eine weitere Gruppe von Verfahren zur Genotypisierung basiert auf der Ermöglichung oder der Unterbindung von Amplifikationen durch die Hybridisierung von Oligonukleotiden un-

ter stringenten Bedingungen. So werden bei der PCR-SSP, einer PCR mit sequenz-spezifischen Primern, die auch als ASA (*allele specific amplification*) oder ARMS (*amplification refractory mutation system*) bezeichnet wird, selektive Primer verwendet, die an der polymorphen Stelle mit der Templat-DNA hybridisieren. Die Hybridisierungsbedingungen werden dann entweder so stringent gewählt, dass der Primer nur mit der vollständig komplementären Sequenz hybridisiert, oder die Primersequenz wird so gewählt, dass das 3'-Ende des Primers auf der polymorphen Sequenzposition zu liegen kommt. In beiden Fällen erfolgt eine Amplifikation idealerweise nur dann, wenn die Primerbindungsstelle im Templat perfekt komplementär zur Primersequenz ist.[92,93]

Die Amplifikation eines Allels bestimmter Sequenz kann auch durch die Verwendung von PNAs (*peptide nucleic acids*) gesteuert werden, deren Sequenz zu der Zielsequenz zwischen den Primerbindungsstellen komplementär ist. DNA/PNA-Hybride sind sehr stabil und verhindern bei ihrer Ausbildung die effiziente Verlängerung der Primer und damit die Amplifikation des PCR-Produkts. Eine Basenfehlpaarung im DNA/PNA-Hybrid destabilisiert diesen jedoch derart, dass sich die PNA-Sonde leicht ablöst oder unter den Reaktionsbedingungen gar nicht erst bindet und so die Blockade aufgehoben ist, so dass es zu einer Amplifikation kommen kann. Dadurch lassen sich Sequenzvarianten bei der Amplifikation diskriminieren.[94,95]

Die Entwicklung von RT-PCR-Methoden eröffnete in den letzten Jahren auch neue Möglichkeiten zur Genotypisierung, die insbesondere den Nachweis von SNPs vereinfachen. Hierbei wird entweder die Generierung des amplifikationsbedingten Fluoreszenzsignals von der polymorphen Sequenz beeinflusst („PCR-gekoppelte Analyse“) oder es wird im Anschluss an die PCR eine Schmelzkurve der PCR-Produkte aufgezeichnet, anhand derer sich die Polymorphismen identifizieren lassen („Schmelzkurven-Analyse“). So lassen sich für eine PCR-gekoppelte Analyse z. B. TaqMan-Sonden und *Molecular Beacons* derart konstruieren, dass sie bevorzugt oder – unter entsprechend stringenten Reaktionsbedingungen – ausschließlich mit der perfekt komplementären Sequenz hybridisieren und damit eine Signalgenerierung unterbunden wird, wenn innerhalb dieses Abschnittes auf der Templat-DNA Sequenzvariationen vorliegen.[43,96] Bei der Schmelzkurvenanalyse hingegen wird die durch Basenfehlpaarungen bedingte Veränderung der Schmelztemperatur von dsDNA-Fragmenten gemessen. Im einfachsten Falle kann dazu ein dsDNA-sensitiver Farbstoff wie SYBR-Green I verwendet werden, womit sich kleine Sequenzvariationen wie SNPs allerdings nur sehr schwer nachweisen lassen.[97-99] Diese lassen sich jedoch sehr einfach anhand des Schmelzverhaltens sequenzspezifischer Sonden wie den *Molecular Beacons* oder den Hybridisierungssonden nachweisen.[42,100-103] Der Vorteil von Schmelzkurvenanalysen gegenüber den PCR-gekoppelten Analysen liegt in ihrem Potenzial, mehrere im Reaktionsansatz vorliegende Sequenzvariationen simultan nachweisen zu können.[104]

1.4 DNA-analytische Verfahren zur Chimärenanalytik

Wenn sich alles gleicht, zählt der Unterschied.

Kurt Brugger

Die Chimärenanalytik erlaubt quantitative Aussagen über die Mischungsverhältnisse von Proben unterschiedlicher (klonaler) Herkunft. Neben einigen immunologischen Methoden wie der immunohistochemischen Untersuchung von MHC-I und AB0-Blutgruppen-Antigenen[105] oder durchflußzytometrischen Untersuchungen[106] kommen vor allem PCR-basierte Methoden zum Einsatz, welche individuelle Sequenzvariationen quantifizieren. Diese Methoden sind also im Prinzip quantitative Genotypisierungen.[107-110]

Die Chimärenanalytik ist von kontinuierlich wachsender Bedeutung in der molekularen Biologie im Allgemeinen sowie der molekularen Medizin im Besonderen. Hier findet sie z. B. Anwendung bei populationsbiologischen Fragestellungen[111] in der Analyse der post-operativen Entwicklung nach der Transplantation von hämatopoetischen Stammzellen[108,110] oder solider Organe[112-114] sowie in der Verfolgung des Wachstums definierter Zellklone in einem Tumor.[77] Während die Analyse des klonalen Wachstums in Tumoren noch im Bereich der Grundlagenforschung anzusiedeln ist, kommt der Chimärenanalyse nach Organtransplantationen in mehrfacher Hinsicht erhebliche diagnostische Bedeutung zu.

Die Transplantation hämatopoetischer Stammzellen von einem gesunden Spender in einen leukämiekranken Empfänger soll dort das ursprüngliche, tumoröse blutbildende System ersetzen. War die Zerstörung des eigenen Systems durch Chemo- oder Strahlentherapie nicht vollständig, so besitzt der Empfänger nach der Transplantation einen gemischten hämatopoetischen Chimärismus (MHC, *mixed hematopoietic chimerism*). Die Frage, ob der MHC als Auslöser von Rückfällen in Frage kommt, ist noch nicht entschieden und Gegenstand aktueller Forschung, in welcher eine sensitive Chimärenanalytik zur Anwendung kommt.[114,115] Ebenso wie nach einer Knochenmark-Transplantation kann es auch nach der Transplantation solider Organe wie Herz, Leber, Lunge und Niere zu Folgeerkrankungen kommen.[116,117] Spender-Stammzellen des Immunsystems können das Transplantat verlassen und im Blut über Jahre hinweg weiterexistieren.[118] Diese immunkompetenten Zellen können im Empfänger Immunreaktionen vermitteln und gegen den Empfänger gerichtete zytotoxische T-Zellen und Antikörper bilden. Diese als *graft-versus-host disease* (GVHD) bezeichnete Erkrankung tritt auch dann auf, wenn das Immunsystem des Empfängers durch Immunsuppressiva unterdrückt wird, um Abstoßungsreaktionen gegen das transplantierte Organ zu vermeiden.[113,119] Der Nachweis und die Verfolgung der Entwicklung dieses peripheren Mikrochimärismus ist essentiell für eine frühe Aussage über die Gefahr einer auftretenden GVHD.[120] Im umgekehrten Fall, dem *intragraft*-Mikrochimärismus, wandern Zellen des Empfängers in das Spenderorgan ein. Das Ausmaß und die Entwicklung der Einwanderung kann durch die Chimärenanalytik von Biopsieproben aufgeklärt werden.[110,112]

Entsprechend der großen Bedeutung der Chimärenanalytik für die molekularmedizinische Diagnostik gibt es eine Reihe unterschiedlicher Analysesysteme, von denen im folgenden

die prominentesten Verfahren, die auf DNA-analytischen Methoden beruhen, kurz vorgestellt werden.

Nukleinsäureanalytische Methoden zur Chimärenanalytik beruhen meist auf der selektiven Amplifikation spezifischer DNA-Sequenzen durch die PCR. Dabei werden entweder polymorphe Sequenzen unterschieden oder solche Sequenzen nachgewiesen, die nur in einem Teil der Mischprobe vorkommen. So lässt sich beispielsweise mit einer für Y-Chromosomen spezifischen PCR sehr einfach ein Geschlechtschimärismus analysieren, wie er z. B. nach einer *sex-mismatched allogenic blood stem cell transplantation* vorliegt.[120-122] Es gibt bereits Real-time-Verfahren zur quantitativen Y-Chromosomen-spezifischen PCR (QYCS-PCR). Damit lässt sich eine männliche Zelle selbst unter 100 000 weiblichen Zellen nachweisen.[123] Allerdings ist die Methode auf die spezielle Situation eines Geschlechtschimärismus beschränkt, und die Sensitivität ist nur nutzbar, wenn der Proben-Anteil männlicher Herkunft gering ist.

Als geschlechtsunabhängige Zielsequenzen zur Chimärenanalytik dienen die schon im Abschnitt **1.3** erwähnten RFLPs[122] und STRs,[124-127] ebenso wie eine *tandem-repeat*-Klasse mit längeren Grundeinheiten als bei STRs, die *variable number of tandem repeats* (VNTRs).[107,128-130] Im Unterschied zur reinen Genotypisierung wird hier das Verhältnis unterschiedlicher Allele in einer Mischprobe quantifiziert. Dies geschieht meistens durch densitometrische Auswertungen der elektrophoretisch getrennten PCR-Fragmente. Eine Sequenz mit weniger als 1-5 % Anteil an der Mischprobe kann von diesen Verfahren nicht mehr nachgewiesen werden.[128] Die Gründe dafür liegen zum einen im relativ geringen dynamischen Bereich densitometrischer Bestimmungen, zum anderen in dem Umstand, dass kürzere Sequenzen in der PCR präferenziell amplifiziert werden. Wenn die Zielsequenzen des kürzeren allel-spezifischen Fragments in Überzahl vorliegen, kann es zu einem sogenannten *allelic drop-out* kommen, wodurch die Zielsequenzen des längeren Fragments nicht mehr nachgewiesen werden können.[131]

1.5 Molekulare Grundlagen von Krebserkrankungen

Sane sicut lux seipsam et tenebras manifestat, sic veritas norma sui et falsi est.
(Wahrlich gleich wie das Licht sich selber und die Dunkelheit offenbart,
so ist die Wahrheit das Maß ihrer selbst und des Irrtums)

Spinoza in Ethik

Nach heutigem Verständnis sind Tumore das Ergebnis genetischer Veränderungen, die sich im Rahmen eines Mikroevolutionsprozesses in einzelnen Zellen anreichern und die Störung grundlegender zellulärer Prozesse, insbesondere die der Zellteilungsregulation, bedingen. Durch den Ausfall von DNA-Reparatursystemen, sowie Störungen der DNA-Synthese und von Kontrollsystemen der Segregation können sich nach den Zellteilungen der Tumorzellen umfangreiche genetische Veränderungen ansammeln, darunter Sequenzänderungen sowie

Translokationen, Deletionen und Amplifikationen von Genen, Chromosomenabschnitten oder ganzen Chromosomen. Die involvierten Gene definieren die Fehlsteuerung in den Zellteilungsprozessen, bei der Kontaktinhibition sowie des Migrationsverhaltens und bestimmen damit die Charakteristik, Aggressivität und Therapierbarkeit von Tumoren. Bereits heute werden die Erkenntnisse der molekularen Medizin über solche Veränderungen für Diagnose, Prognose sowie Design patientenspezifischer Therapien genutzt. Unverzichtbare Voraussetzung ist hierzu allerdings, dass die eingesetzten analytischen Methoden hohe Sensitivität, Präzision und Zuverlässigkeit aufweisen.

Bei der Transformation einer normalen Zelle in eine Tumorzelle sind genetische Elemente betroffen, welche das Zellwachstum, die Zellproliferation oder die Apoptose kontrollieren. Die Gene lassen sich dabei in zwei Gruppen einteilen: die (Proto-) Onkogene und die Tumorsuppressorgene. Während (Proto-)Onkogene positive Regulatoren von Wachstum und Proliferation darstellen, handelt es sich bei den Tumorsuppressorgenen um die negativen Regulatoren, welche die Wachstums- und Proliferationsprozesse stoppen oder die Apoptose initiieren. Der von vielen inneren und äußeren Faktoren beeinflusste Wachstums- und Teilungszyklus der Zellen wird so durch das Zusammenspiel positiver und negativer Regulatoren kontrolliert. Die Tumortransformation kann durch eine fehlerhafte Aktivität bzw. einen Funktionsgewinn („*gain of function*“) eines (Proto-)Onkogens bedingt sein, welches dann als „Onkogen“ bezeichnet wird oder durch den Funktionsverlust („*loss of function*“) eines Tumorsuppressorgens.

1.5.1 Onkogene

(Proto-)Onkogene kodieren meist für Proteine in Signaltransduktionswegen, deren Aktivierung letztlich die Zellproliferation anregen. Die Signalketten beginnen dabei mit den Interaktionen von Wachstumsfaktoren mit Rezeptoren auf der Zellmembran und führen über Signalkaskaden im Zytoplasma schließlich in den Zellkern, wo genetische Programme aktiviert werden, welche die Einleitung von Zellteilungszyklen auslösen. Während der Transformation einer gesunden in eine neoplastische Zelle wird durch Mutationen die Aktivität oder Expression von (Proto-)Onkogenen gesteigert (*gain of function*). Diese überaktiven Produkte der jetzt als Onkogen bezeichneten Gene fördern so die unkontrollierte Zellteilung. Beispiele für (Proto-)Onkogene sind Wachstumsfaktoren (z. B. *sis*, *hst*), Rezeptor-Tyrosin-Proteinkinasen (z. B. *EGFR*, *HER-2/neu*), membrangebundene G-Proteine (z. B. *Ras*, *LCK*), Transkriptionsfaktoren (z. B. *myc*, *fos*, *jun*) und die Gene für Cyclin-D-verwandte Proteine, welche den Zellzyklus regulieren.[132]

Es gibt viele mögliche Ursachen für den Funktionsgewinn von (Proto-)Onkogenen: Eine Sequenzmutation innerhalb oder den Verlust einer regulatorischen, normalerweise inhibitorischen Domäne, die durch eine Translokation verursachte Kombination mit einer aktivierenden Domäne oder die Überexpression. Der Grund für eine Überexpression kann dabei entweder die veränderte Transkriptionsaktivität oder die Amplifikation des Gens selbst bzw. eine Kombination aus beiden sein.

Ein Beispiel eines durch Amplifikation aktivierten Onkogens ist *HER-2/neu*, welches für eine Rezeptor-Tyrosin-Proteinkinase kodiert. Das zum epidermalen Wachstumsfaktorrezeptor *EGFR* strukturell homologe, transmembranöse Glykoprotein wird in vielen Zellen exprimiert und ist involviert in einem ganzen Netzwerk von Signalketten. Mithin spielt *HER-2/neu* eine wichtige Rolle bei Zellwachstum und Entwicklung.[133] Die Aktivitätssteigerung des Rezeptors in Tumoren ist in vielen Fällen Folge einer stark gesteigerten Expression. Normalerweise dimerisiert der Rezeptor nach Ligandenbindung, wodurch die zytosolische Tyrosinkinase-domäne aktiviert wird. Die Überexpression führt zu einer höheren Dichte des Rezeptors in der Zellmembran, so dass es vermehrt auch ohne Ligandenbindung zur Dimerisierung und Aktivierung kommt.[134] Ursache der Überexpression ist meist eine Genamplifikation von *HER-2/neu* im Genom, was zu einer Erhöhung der Transkriptmenge und auf Proteinebene schließlich zu einer Erhöhung der Rezeptordichte in der Zellmembran führt.[135]

HER-2/neu wird in 20-30 % aller Brustkrebsfälle überexprimiert,[136] was mit einer erhöhten Resistenz der Tumore gegen Chemotherapien einhergeht.[137] In diesem Falle kann das Tumorwachstum durch eine Applikation von Herceptin, einem gegen den Rezeptor gerichteten, humanisierten monoklonalen Antikörper, gestoppt werden.[138,139] Der *HER-2/neu*-Status ist daher von Bedeutung für die Beurteilung, welche Behandlung für die jeweilige Patientin den größten Heilungserfolg verspricht.

Die verbreitetsten Methoden zur Detektion einer HER-2-Überexpression beruhen auf dem immunohistochemischen Nachweis des Rezeptors, z. B. mittels HERCEPTest,[140] sowie dem Nachweis einer Genamplifikation durch Fluoreszenz-*in-situ*-Hybridisierung (FISH).[141] Beide Techniken sind von der U.S. Food and Drug Administration (FDA) für die klinische Diagnostik zugelassen. Neuere Methoden basieren auf PCR-Techniken zum quantitativen Nachweis der Gensequenz oder der Transkripte.[142-148] Die Quantifizierung der genomischen *HER-2/neu*-Kopienzahl hat gegenüber der Bestimmung der Transkriptmengen den Vorteil einfacher Probenvorbereitung aufgrund der höheren Stabilität von DNA im Vergleich zu RNA, erfordert jedoch präzisere Nachweismethoden, da die relativen Unterschiede in der Genamplifikation meist geringer sind als die daraus resultierenden relativen Unterschiede in den Transkriptmengen zwischen Normal- und Tumorgewebe.

1.5.2 Tumorsuppressorgene

Tumorsuppressorgene kodieren für Produkte, welche die Zellproliferation inhibieren. Ihr Funktionsausfall, z. B. in Folge einer Mutation oder des Genverlusts, begünstigt die unkontrollierte Zellteilung. Phänotypisch relevant ist meistens nur der Funktionsverlust beider Allele einer Zelle, da andernfalls das verbleibende, intakte Allel die regulatorische Funktion aufrechterhalten kann. Der Funktionsverlust kann verursacht sein durch Mutationen sowie durch Allelverluste, letztere infolge von Segregationsfehlern, Deletionen, Rekombinationen oder Genkonversionen.

Die Bedeutung negativ regulatorischer Gene bei der Krebsentstehung wurde erstmals bereits Ende der 1960er Jahre am *Xiphophorus*-Tiermodell entdeckt.[149] Das erste humane Tumorsuppressorgen, *RB1*, wurde schließlich 1971 von KNUDSON beschrieben.[150] Es ist an

der Ausbildung von Retinoblastomen beteiligt, die in Familien mit familiärer Belastung, d. h. einer erblich bedingten Prädisposition zur Ausbildung von Retinoblastomen, sehr häufig und schon im frühen Kindesalter auftreten, wobei oft multifokale Tumore auftreten. Die sporadische Form ist weitaus seltener (etwa 1:12 000)[151] und tritt meist unifokal auf. Während somatische Mutationen in einer intakten Kopie des *RB1*-Gens bei gesunden Menschen rezessiv sind, haben sie bei familiär belasteten Menschen eine „dominante" Wirkung.

Auf Grund dieser Befunde formulierten KNUDSON das Zwei-Treffer-Modell zur Inaktivierung von Tumorsuppressorgenen. Familiär belastete Menschen erben nur eine intakte Kopie des Gens, da die andere Kopie in der Keimbahn durch eine Mutation funktionslos gemacht oder verloren wurde. Nun reicht während der Ontogenese der Ausfall der einzig noch intakten Kopie des Gens in einer beliebigen Zelle der Netzhaut aus, damit sich ein Retinoblastom entwickeln kann. Demgegenüber ist es in einer Zelle mit zwei intakten *RB1*-Kopien vergleichsweise unwahrscheinlich, dass infolge spontaner Mutationsereignisse beide *RB1*-Kopien ausfallen.

Durch zell- und molekularbiologische Untersuchungen wurden sechs verschiedene mechanistische Wege nachgewiesen, über die phänotypisch unauffällige Zellen mit einem durch Mutation inaktiviertem Allel eines Tumorsuppressorgens das funktionale Allel mit der Wildtypsequenz verlieren können:[152] (i) Verlust des Wildtyp-Chromosoms durch *non-disjunction*, (ii) Verlust des Wildtyp-Chromosoms durch *non-disjunction* mit anschließender Duplikation des mutierten Chromosoms, Ersatz des Wildtypallels durch das mutierte Allel über (iii) mitotische Rekombination oder (iv) Genkonversion, (v) Verlust des Wildtypallels durch Deletion und (vi) Inaktivierung des Wildtypallels durch Punktmutation.

Bis auf den letzten Punkt lässt sich der Verlust des Wildtypallels durch die Analyse polymorpher DNA-Marker wie den RFLPs, VNTRs und STRs nachweisen, die innerhalb bzw. in der Nähe des untersuchten Gens liegen und für welche das betreffende Individuum heterozygot ist. Der Verlust eines Gens, der auch den betrachteten Polymorphismus einschließt, bewirkt damit in den meisten Fällen auch den Verlust der Heterozygosität. Durch LOH (*loss of heterozygosity*)-Analysen wurden und werden noch immer neue Loci von Tumorsuppressorgenen kartiert und identifiziert.[153]

Die Wirkungsweisen der Tumorsuppressorgene sind sehr heterogen. Sie können an der Kontrolle des Zellwachstums und der Zellteilung (z. B. *RB*, *p16*, *p53*), der DNA-Reparatur bzw. dem Erhalt der Genomintegrität (z. B. *hMLH1*, *hMSH2*, *BRCA1*, *p53*), der Kontrolle der Apoptose (z. B. *p53*) oder der Inhibition von Wachstums- und Transkriptionsfaktoren (z. B. *NF1*) beteiligt sein. Außerdem finden sich unter den Tumorsuppressorgenen u. a. sowohl Proteinkinasen (z. B. *MAP2K4*) wie auch Proteinkinase-Inhibitoren (z. B. *CDKN1C*, *CDKN2A*).[132]

Das am häufigsten in Tumoren deletierte oder mutierte Tumorsuppressorgen ist *p53*.[154-157] Nachgewiesen wurde sein Genprodukt erstmals 1979 in SV40-infizierten und neoplastischen Zellen, in denen es mit dem SV40 *Large T-Antigen* komplexiert vorlag, weshalb man zunächst davon ausging, dass es sich hierbei um ein Tumorantigen handele.[158-160] Später

stellte sich jedoch heraus, dass sein Genprodukt, p53, die transformierende Wirkung vieler Onkogene supprimiert[161] weshalb es als Tumorsuppressor eingestuft wurde, wenngleich auch bestimmte mutierte Varianten onkogene Wirkungen zeigen können.[132]

Der Funktionsverlust von p53 bewirkt eine drastisch gesteigerte genetische Instabilität (in *p53*$^{-/-}$-Zellen ist die Wahrscheinlichkeit von Onkogenamplifikationen im Vergleich zur Wildtypzelle um das Millionenfache erhöht[162]) sowie den Verlust der Fähigkeit, bei DNA-Schädigungen das Zellwachstum und die Zellteilung zu stoppen[163,164] bzw. die Apoptose einzuleiten.[165,166] Das p53-Protein kontrolliert den Zellzyklus-Kontrollpunkt am Übergang von der G_0/G_1- in die S-Phase,[167] der für den Erhalt der Genomintegrität wesentlich ist.[168,169] Wegen der vielfältigen Kontrollfunktionen von p53 im Zellzyklus, bei der DNA-Replikation und -Reparatur sowie der Apoptose ist anzunehmen, dass der Aktivitätsstatus von p53 in einem spezifischen Tumor einen erheblichen Einfluss auf die weitere Entwicklung des Tumors hat.

Übliche Therapien von Tumorerkrankungen basieren oftmals auf der irreparablen Schädigung der DNA teilungsaktiver Zellen, hauptsächlich durch Doppelstrangbrüche, ausgelöst durch ionisierende Strahlung oder DNA-schädigende Chemikalien. Die heilende Wirkung durch diese Therapieformen beruht auf der Tatsache, dass sich teilungsaktive, geschädigte Zellen durch Apoptose selbst eliminieren. Ein wichtiger Schritt in der Auslösung des Apoptoseprogramms nach DNA-Schädigung ist jedoch die Aktivierung von p53, der in p53-defizienten bzw. p53-inaktiven Zellen nicht erfolgen kann.[170,171] Diese Zellen können teilweise trotz größter genetischer Schäden überleben und sich weiter teilen. Durch die Therapie unterliegen sie tumorspezifischen *in-vivo*-Selektionsprozessen,[172] wodurch vermutlich die Entwicklung noch aggressiverer Subklone begünstigt wird. Inzwischen ist durch zahlreiche Publikationen nachgewiesen, dass die Inaktivierung von p53 für unterschiedliche Tumore mit einer deutlich schlechteren Prognose bzw. einer reduzierten Wirksamkeit chemotherapeutischer Behandlungen einhergeht.[z. B. 173-177]

Die Diagnose des p53-Status ist damit ein wichtiges Kriterium zur Beurteilung der Prognose sowie zur Wahl eines adäquaten Therapiekonzepts, da Strahlen- und Chemotherapien bei p53-inaktiven Tumoren von verminderter Wirksamkeit sind und wahrscheinlich noch die Evolution aggressiverer Tumore fördern.

1.5.3 Mobile genetische Elemente

Mobile genetische Elemente sind DNA-Abschnitte, welche z. B. durch Rekombinationsprozesse an andere Stellen im Genom oder in ein anderes Genom versetzt (transponiert) werden können. Transpositionen sind häufig verbunden mit auftretenden Deletionen oder Inversionen. Mobile genetische Elemente sind auf vielfältige Weise in der Lage ein Genom zu beeinflussen und haben daher entscheidenden Einfluss auf die Evolution,[178] können mitunter aber auch Krankheiten, so z. B. Krebs, hervorrufen.[179]

Erstmals postuliert wurde die Existenz solcher mobiler Elemente auf Grund genetischer Studien am Mais 1951 von B. MCCLINTOCK.[180,181] Ihr Konzept eines dynamischen Genoms

wurde jedoch erst in den 1970er Jahren allgemein akzeptiert.[182] Mittlerweile ist eine große Anzahl mobiler genetischer Elemente in Phagen, Bakterien, Pilzen, höheren Pflanzen, Vertebraten und Insekten beschrieben worden. Die mobilen Elemente können entsprechend dem Mechanismus der Transposition in zwei Klassen eingeteilt werden:

- Transposons. Die Transposition erfolgt über ein DNA-Intermediat, initiiert durch die Spaltung der Wirts-DNA durch das Enzym Transposase. Bei der direkten Transposition wird das mobile Element aus der DNA ausgeschnitten und an anderer Stelle wieder eingesetzt. Beim replikativen Weg wird das Element repliziert. Das Original bleibt an der ursprünglichen Stelle, während die Kopie durch Rekombinationsprozesse an einer anderen Stelle eingebaut wird.[183,184] Den einfachsten Aufbau zeigen die so genannten Insertionssequenzen (IS) der Prokaryoten. Die 800-2000 Bp-langen IS-Elemente bestehen nur aus einem Transposasegen, welches beidseitig von invertierten Wiederholungssequenzen eingerahmt ist.[185,186] Zusammengesetzte Transposons (Tn) bestehen aus zwei IS-Elementen, die ein Gen, meistens ein Antibiotika-Resistenzgen, flankieren.[187]

Beispiele für Transposons sind die von McClintock entdeckten Ac- und Ds-Elemente im Maisgenom,[180] die in ein- und mehrzelligen Eukaryoten weit verbreitete Familie der tc1- bzw. mariner-line elements (TLE bzw. MLE), die z. B. im Genom des Menschen in einigen tausend Kopien vorliegen,[188] sowie die P-Elemente im Drosophila-Genom, die experimentell zur Transformation von Drosophila-Zellen verwendet werden.[189,190]

- Retroelemente. Im Gegensatz zu den Transposons erfolgt hier die Transposition über ein RNA-Intermediat, welches durch eine Reverse Transkriptase in DNA umgeschrieben wird.[178] Retroelemente kommen in allen eukaryotischen Genomen vor. Sie lassen sich klassifizieren in Pseudo- und Retrogene, Retroposons, Retrotransposons und Retroviren:[191]

Pseudogene sind die Folge der reversen Transkription und Reintegration eines zellulären mRNA-Moleküls. Transkribierte Pseudogene unter Kontrolle eines Promotors bezeichnet man als Retrogene, wie z. B. die SINE (*short interspersed nuclear elements*)-Sequenzen der Säuger,[192] zu denen auch die humane Alu-Repeat-Familie gehört, die 3-6 % des humanen Genoms ausmacht.[193]

Retroposons besitzen einen internen Promotor und ein oder mehrere offene Leseraster, flankiert von speziellen repetitiven Sequenzen, den so genannten LTRs (*long terminal repeats*) mit regulatorischen Sequenzen.[194] Die als LINE (*long interspersed nuclear elements*)-Sequenzen bezeichneten Retrotransposons der Säuger kodieren für eine reverse Trankriptase [195] und werden in Tumoren verstärkt kopiert.[196,197]

Retroviren enthalten RNA, die nach der Infektion einer Wirtszelle durch die Reverse Transkriptase in DNA umgeschrieben wird, welche als Provirus in das Genom der Wirtszelle integriert werden kann.[198] Aus dem Provirus kann durch Transkription wieder die virale RNA sowie die mRNA zur Translation viraler Proteine gebildet werden.[199] Viele provirale Sequenzen sind jedoch durch Mutationen mehr oder weniger stark degeneriert, so dass sie die Zelle nicht mehr verlassen, mitunter aber noch in das Wirtsgenom integriert werden können.[200] Die Grenzen zwischen endogenen Retroviren und Retrotransposons sind flie-

ßend. Manche Viren besitzen ein eigenes Onkogen (v-onc), das auf ein Proto-Onkogen des Wirtsorganismus zurückgeht, welches in das Virusgenom integriert wurde und dort Sequenzänderungen akkumuliert hat.

1.6 Elektronische Datenverarbeitung

Irren ist menschlich. Aber wenn man richtig Fehler machen will, braucht man Computer.

Harald Lesch

Durch die schnell zunehmende Verbreitung automatisierter Messsysteme in der Biologie und Medizin, nicht zuletzt durch die Entwicklung immer neuer Hochdurchsatz-Verfahren, steigt das Datenaufkommen überproportional an und mithin wächst der Bedarf an Systemen zur effizienten Verwaltung und Auswertung der Daten. Dies macht den Einsatz elektronischer oder Computer-gestützter Auswertungsverfahren zwingend notwendig.

Messgrößen sind physikalische Größen, denen eine Messung gilt. Die Messung liefert einen Schätzwert für den wahren Wert der Messgröße, das Messergebnis. Unter einem Wert versteht man dabei das Produkt aus einer reellen Zahl und der Einheit einer physikalischen Größe. Messunsicherheiten entstehen durch statistische und systematische Abweichungen der Messergebnisse vom wahren Wert (siehe auch **7.1**). Ursachen für statistische Abweichungen sind das Auflösungsvermögen des Messinstruments und nichtkonstante, unbestimmbare Umwelteinflüsse auf die Messung. Dadurch bedingte Messfehler lassen sich durch die Mittelung von Ergebnissen wiederholter Messungen oder elektronischer Verfahren zur Rauschfilterung minimieren. Systematische Abweichungen entstehen u. a. durch falsche Kalibrierung oder konstruktive Mängel der Messanordnung oder durch eine Einflussnahme des Messvorgangs auf die Ermittlung des Messwerts. Diese Art der Abweichung lässt sich häufig durch geeignete mathematische Verfahren eliminieren.

Für eine elektronische Verarbeitung müssen die in den meisten Fällen in analoger Form (d. h. wert- und zeitkontinuierlich, z. B. Temperatur, Gewicht, Stromstärke) anfallenden Messdaten zunächst durch geeignete physikalische Instrumentierungen in einen analogen Spannungswert gewandelt werden, der dann digitalisiert (d. h. in ein werte- und zeitdiskretes System transformiert) wird. Die meisten modernen Messsysteme (z. B. Spektrophotometer, Geldokumentationsanlagen, ELISA-*Reader*, BIAcore, EKG-/EEG-Apparate, *Phosphoimager*, *Laserscanning*- und Raster- sowie Tunnelelektronen-Mikroskope, RT-PCR-Geräte usw.) zeichnen die Daten schon in elektronischer Form auf, so dass die elektronische Verarbeitung der Daten obligatorisch ist. Der Rechner kann die Daten über die digitale Schnittstelle abfragen und speichern bzw. auswerten. Die Abfrage und Auswertung wird von der Software gesteuert.

Bei der *on-line*-Verarbeitung werden die digitalisierten Daten von der Software direkt weiterverarbeitet. Reine *on-line*-Verarbeitungen sind selten und lohnen nur bei Langzeitmessungen mit niedriger Frequenz der Datenaufzeichnung (*sample*-Rate) oder bei Aufgaben-

stellungen, bei denen der Anwender aufgrund aktueller Messergebnisse in den Prozessablauf eingreifen muss. Bei der *off-line*-Verarbeitung werden die Rohdaten zunächst aufgenommen und gespeichert. Nach Abschluss der Messung wird dem Anwender durch die Software Gelegenheit gegeben, die Daten in gewünschter Form aufzuarbeiten.

Die in den letzten Jahren drastisch gestiegene Leistungsfähigkeit der EDV-Anlagen ermöglichte die Entwicklung immer anwenderfreundlicherer Software, so dass zunehmend auch „Nicht-Spezialisten" in der Lage sind, komplexe Auswertungen großer Datenmengen durchzuführen. Dennoch fordern auch heute noch viele Anwendungen in der Mess- und Regeltechnik vom Anwender mehr EDV-Kenntnisse als notwendig. Dies liegt nicht zuletzt daran, dass die Erstellung professioneller Programme zur Automatisierung von Messverfahren und Auswertungen zeit- und kostenintensiv ist.

Dabei trägt die Qualität der Software entscheidend zur Leistungsfähigkeit des Messsystems bei. Auswerteprogramme, die schwer zu bedienen und unübersichtlich sind, erfordern im günstigsten Falle unnötig hohe Investitionen an Zeit und zudem fachkundiges Personal, schlimmstenfalls sind sie sogar mitverantwortlich für fehlerhafte Interpretationen der Daten durch den Anwender. Gute Programme stehen hier in der Verantwortung, durch eine übersichtliche und geschickte Benutzerführung und Präsentation der Daten Fehler bei der Auswertung vermeiden zu helfen sowie dem Anwender nicht nur Ergebnisse zu präsentieren, sondern ihm auch bei der Wertung der Daten zu assistieren.

1.7 Softwareentwicklung

640 kB sollten genug für jeden sein.

Bill Gates, 1981

Der Begriff Software beschreibt Programme der elektronischen Datenverarbeitung. Dabei kann es sich um Betriebssysteme, Dienstprogramme oder Arbeitsprogramme handeln. Als Programm bezeichnet man die zur Lösung eines Problems oder Aufgabenkomplexes erforderliche Folge von Befehlen, die von einem Computer ausgeführt werden können. Die Programmierung umfasst die Abstraktion einer Problemstellung zu einem Fachmodell, welches gegebenenfalls zu einem mathematischen Modell präzisiert wird. Das zur Lösung des Problems geeignete Verfahren wird als Algorithmus bezeichnet. Das oder die Lösungsverfahren werden schließlich in direkt oder indirekt computerverständlicher Form kodiert. Zur Kodierung werden in den meisten Fällen höhere Programmiersprachen verwendet, bei denen es sich ihrerseits um Software handelt, welche einen menschenverständlichen Kode, den sog. Quellkode, in eine direkt vom Rechenwerk verwertbare Form, den binären Maschinenkode übersetzt.

Die zur Entwicklung der Software in dieser Arbeit verwendete Sprache ist Pascal, die 1972 an der Eidgenössischen Hochschule Zürich von NIKOLAUS WIRTH zu Lehrzwecken entwickelt und nach dem französischen Mathematiker BLAISE PASCAL (1623 - 1662) benannt wurde.

Bei Pascal handelt es sich um eine *Compiler*-Sprache, bei welcher der Quellkode nach Abschluss der Kodierung vollständig in Maschinensprache übersetzt, d. h. kompiliert wird. Das Kompilat ist dann die ausführbare Programm-Datei als Ergebnis der Softwareentwicklung.

Ein Grundkonzept der Sprache Pascal ist die *Top-down*-Programmierung, bei welcher das Programmierprojekt in immer kleinere Programmeinheiten aufgespalten wird, die über definierte Schnittstellen miteinander verbunden sind. Dieser Vorgang wird soweit fortgesetzt, bis kleine direkt kodierbare Einheiten entstehen. Diese Einheiten werden als definierte, eigenständige Unterprogramme programmiert, welche die gleiche prinzipielle Struktur wie das Hauptprogramm aufweisen und geschachtelt werden können. Die Strukturiertheit bezieht sich auch auf die Datenverwaltung, die eine Definition von Datentypen erlaubt und eine strikte Deklaration von Variablen bestimmten Datentyps erfordert. Unterprogramme können ihrerseits wieder in Bibliotheken, den so genannten `Units`, zusammengefasst werden und so in einfacher Weise von anderen `Units` oder Projekten angesprochen werden, indem diese die betreffende `Unit` einbinden. Durch diese Konzepte ist Pascal zu einer Galionsfigur der strukturierten Programmierung geworden. Dennoch wurde seither die Programmiersprache ständig weiterentwickelt.

Schon in den siebziger Jahren des 20. Jahrhunderts wurden erste Versuche unternommen, neue Methoden des Dialogs zwischen Mensch und Maschine zu entwickeln. Einer sehr menschlichen Sichtweise der Welt, die Umgebung in definierte Objekte mit ihren eigenen Eigenschaften und Möglichkeiten einzuteilen und anzusprechen, wurde im Konzept einer objektorientierten Programmierung (OOP) Rechnung getragen. Hierbei werden Daten mit den Programmteilen in so genannten Objekten verbunden, welche diese Daten verwalten (kapseln). Objekte können ihre Daten „nach außen“ verbergen oder über öffentliche Schnittstellen, so genannte Eigenschaften, ansprechbar machen. Die Objekte definieren dabei die Möglichkeiten der externen Manipulation ihrer Daten. Aus bestehenden Objekten können neue Objekte abgeleitet werden, welche alle Eigenschaften ihres Vorgängers erben und um weitere Eigenschaften ergänzt werden können. Dazu benötigt der Entwickler nicht einmal Einblick in die Implementation des Vorgängerobjekts. Objekt-Definitionen, die so genannten Klassen, können so hierarchisch aufgebaut werden, was die Kode-Größe reduziert und die Übersichtlichkeit sowie die Wiederverwendbarkeit von Kodierungen erhöht. Die erste standardisierte objektorientierte *Compiler*-Sprache dürfte wohl das 1985 von BJARNE STROUSTRUP entwickelte C^{++} sein. Kurze Zeit später kam mit Turbo Pascal 5.5 auch die erste objektorientierte Pascal-Version auf den Markt.

Die nächste bedeutende Veränderung betraf nicht das eigentliche Sprachkonzept, sondern die Entwicklungsumgebungen zur Programmierung. Das Anfang der neunziger Jahre angebrochene „Windows-Zeitalter“ steigerte die Anforderungen an Programmierkenntnissen enorm und selbst einfachste Applikationen erforderten viele Hundert Zeilen Quelltext mit komplizierten Objektableitungen zur Gestaltung der Programm- bzw. Bedienoberfläche. Microsoft zeigte mit der Entwicklung von Visual Basic erstmals ein Konzept, das den Programmieraufwand deutlich verringerte und eine fast spielerisch einfache Gestaltung von Windows-Programmen erlaubte. Dieses Prinzip der objekt- und ereignisorientierten visuellen Programmierung wurde 1995 mit der Entwicklungsumgebung Delphi auch für Pascal umgesetzt.

1.8 Zielsetzung der Arbeit

Ja, mach nur einen Plan,
Sei nur ein großes Licht!
Und mach dann noch 'nen zweiten Plan.
Gehn tun sie beide nicht.

Berthold Brecht in der Dreigroschenoper

Ziel dieser Arbeit war die Entwicklung RT-PCR-basierter Methoden für die hochgenaue Quantifizierung genetischer Elemente sowie für die qualitative und quantitative Genotypisierung. Dazu sollten zunächst Studien der grundlegenden Parameter und wesentlichen Eigenschaften des zur Verfügung stehenden RT-PCR-Systems durchgeführt und mögliche Schwachstellen bestimmt und eliminiert werden. Weiterhin sollten Verfahren zur automatisierbaren Auswertung der anfallenden Messdaten entwickelt werden. Die entwickelten Algorithmen zur Datenauswertung sollten sodann in einem Softwareprogramm in anwenderfreundlicher Weise implementiert werden. Abschließend galt es, die Leistungsfähigkeit der biochemischen Verfahren und Auswertetechniken an ausgewählten Anwendungsbeispielen mit Relevanz für die molekulare Genetik und Medizin zu demonstrieren.

2 MATERIALIEN UND METHODEN

2.1 Materialien

2.1.1 Geräte

Tab. 2-1 Geräte.

Gerät	Hersteller
Apple Macintosh G3 PowerPC	Apple, San Jose
Fluoreszenz-Spektrophotometer F-4500	Hitachi/Colora, Lorch
Kapillarelektrophorese ABI PRISM 310	Applied Biosystems, Weiterstadt
LightCycler	Roche Diagnostics, Mannheim (s.a. **2.6.2**)
PC mit Intel Pentium III Prozessor	Unicomp, Gießen
PCR-Gerät GeneAmp PCR-System 2400	Applied Biosystems, Weiterstadt
Reinstwasseranlage RRO 90 CN	Seralpur, Ransbach-Baumbach
UV-Spektrometer U3000	Hitachi/Colora, Lorch
UV-Transilluminator	Bachhofer, Reutlingen
Videodokumentationssystem	Intas, Göttingen

2.1.2 Software

Tab. 2-2 Software.

Programm	Hersteller
ABI Prism Collection 1.0.4	Applied Biosystems, Weiterstadt
ABI Prism GeneScan 2.1	Applied Biosystems, Weiterstadt
Delphi 6.0 Standard Edition	Borland/Inprise, Langen
Excel 2000	Microsoft, Verl
LightCycler3 Data Analysis 3.5.28	Roche Diagnostics, Mannheim
LightCycler3 Run 5.10	Roche Diagnostics, Mannheim
Maple 6	Waterloo Maple, Waterloo
Melt94	MICHIKAWA *et al.*[201]
Oligo 5.0	National Biosciences, Plymouth
Photoshop 6.0	Adobe Systems, San Jose
Windows 98	Microsoft, Verl

2.1.3 Datenbanken

Tab. 2-3 Datenbanken. Die Internet-Seiten sind über das hypertext transfer protocol (http://) aufzurufen.

Datenbank	Internet-Adresse
dbSNP	www.ncbi.nlm.nih.gov/SNP/
European Bioinformatics Institute	www.ebi.ac.uk/
Genome Database	www.gdb.org/
HGBASE	hgbase.cgr.ki.se/
Human Gene Mutation Database	www.archive.uwcm.ac.uk/
Human SNP Database	www.genome.wi.mit.edu/SNP/human/
PubMed Medline	www.ncbi.nlm.nih.gov/PubMed/

2.1.4 Chemikalien und Biochemikalien

Tab. 2-4 Chemikalien und Biochemikalien.

Chemikalie	Hersteller
Acrylamid, *ultra pure*	Life Technologies, Eggenstein
Agarose (Ultra Pure)	Life Technologies, Eggenstein
Azorubin	Aldrich, Steinheim
Bisacrylamid, *ultra pure*	Life Technologies, Eggenstein
Bovines Serumalbumin (BSA)	Roche Diagnostics, Mannheim
Desoxynukleotide (dATP, dCTP, dGTP, dTTP)	Roche Diagnostics, Mannheim
EDTA	Merck, Darmstadt
Ethanol absolut	Merck, Darmstadt
Glucose	Merck, Darmstadt
Ethidiumbromid	Serva, Heidelberg
*Hae*III 5x Reaktionspuffer	USB, Cleveland
*Hin*fI 10x Reaktionspuffer	MBI Fermentas, St. Leon-Rot
Hefe-Extrakt	Gibco BRL, Carlsbad
$MgCl_2$	Roche Diagnostics, Mannheim
Pepton	Gibco BRL, Carlsbad
POP4 (4%iges *performance optimized polymer*)	Applied Biosystems, Weiterstadt
Sorbitol	Sigma-Aldrich, Deisenhofen
SYBR-Green I	Roche Diagnostics, Mannheim
Taq DNA-Polymerase 10x Reaktionspuffer	Promega, Mannheim
*Taq*Start Antikörper	CLONTEC Laboratories, Heidelberg
*Taq*Start Verdünnungspuffer	CLONTEC Laboratories, Heidelberg
TEMED	Merck, Darmstadt
TSR (*template suppression reagent*)	Applied Biosystems, Weiterstadt
Xylencyanol	Serva, Heidelberg

2.1.5 Enzyme

Tab. 2-5 Enzyme.

Enzym	Hersteller
Ampli*Taq* DNA-Polymerase, Stoffel-Fragment	Applied Biosystems, Weiterstadt
*Hae*III	USB, Cleveland
*Hin*fI	MBI Fermentas, St. Leon-Rot
Lysozym	Serva, Heidelberg
Proteinase K	Merck, Darmstadt
RNase A, DNase-frei	QIAGEN, Hilden
Taq DNA-Polymerase	Roche Diagnostics, Mannheim

2.1.6 DNA-Standards

Tab. 2-6 DNA-Standards.

DNA-Standard	Hersteller
GeneScan-500-TAMRA	Applied Biosystems, Weiterstadt
Humane genomische DNA	Roche Diagnostics, Mannheim
pTR54×*Hin*fI	selbst; s. unten

Der Fragmentlängenstandard pTR54×*Hin*fI wurde hergestellt wie beschrieben[202] durch vollständige *Hin*fI-Restriktionsspaltung des Plasmids pTR54. Der mit Auftragspuffer auf eine Konzentration von 100 ng/μl eingestellte Standard beinhaltet die in **Tab. 2-7** aufgeführten Fragmente:

Tab. 2-7 Längenstandard pTR54×HinfI für die Polyacrylamid-Gelelektrophorese (2.5.2). Fragmentlängen (Bp) und Gewichtsmengen (ng) je Mikroliter.

Bp	1061	769	719	517	396	338	398	255	221	220	191	154	145	75
ng	19.7	14.8	13.4	9.6	7.4	6.3	5.5	4.7	4.1	4.1	3.5	2.9	2.7	1.4

Tab. 2-8 Längenstandard GeneScan-500-TAMRA für die Kapillarelektrophorese (2.5.4).

Nt	35	50	75	100	139	150	200	250*	300	340*	350	400	490	500

* Diese Fragmente werden bei der Längenberechnung nicht berücksichtigt, da sie ein abnormes Laufverhalten aufweisen.

2.2 Herkunft und Lagerung von Blut- und DNA-Proben

2.2.1 Blutproben

Alle in dieser Arbeit verwendeten Blutproben stammen von nicht miteinander verwandten freiwilligen Probanden mitteleuropäisch-kaukasischen Typs. Das Blut wurde in 10ml-Monovetten (Sarstedt, Nürnbrecht) als K-EDTA-Blut abgenommen. Aus vier Aliquots von je 200 μl wurde sofort die Lymphocyten-DNA isoliert (siehe **2.4.1**), die restliche Blutmenge wurde zur späteren bedarfsgemäßen DNA-Isolation in 200μl-Aliquots bei -20 °C gelagert.

2.2.2 DNA-Proben von Zelllinien

Zellen der humanen Brustkrebs-Zelllinien MDA-468 (ATCC Nr. HTB-132), MCF-7 Cl.18 (ATCC Nr. HTB-22) und SK-BR-3 (ATCC Nr. HTB-30) stammen von der American Type Culture Collection (ATTC). Die schockgefrorenen und bei -70 °C in flüssigem Stickstoff gelagerten Zellproben wurden freundlicherweise von B. BRANDT, Münster, zur Verfügung gestellt. Die DNA wurde wie in **2.4.1** beschrieben isoliert und bei -20 °C gelagert.

2.2.3 DNA-Proben von Tumorpatienten

Tumor- sowie zugehörige Normalgewebe-Proben anonymisierter Patienten mit Blasentumoren wurden freundlicherweise zur Verfügung gestellt von J. SERTH, Medizinische Hochschule Hannover. Die DNA wurde extrahiert wie unter **2.4.1** beschrieben und anschließend bei -20 °C gelagert.

2.2.4 DNA-Proben von Poeciliidae-Spezies

Fische der *Poeciliidae* wurden bezogen vom lokalen Zierfischzüchter (*Girardinus falculatus*; *G. metallicus*; *Heterandria formosa*; *Lebistes reticulatus*; *Priapella compressa*; *Xiphophorus guentheri*; *X. variatus*) sowie von F. ANDERS und A. ANDERS, Institut für Genetik der Justus-Liebig-Universität Gießen (*X. helleri catemaco*; *X. helleri*; *X. maculatus*; *X. variatus*; Hybride von *X. helleri* × *X. maculatus*, über mehr als 50 Generationen rückgekreuzt mit *X. helleri*). Die DNA wurde wie in **2.4.1** beschrieben aus dem Gehirn der in Eiswasser anästhesierten Fische isoliert und dann bis zum Gebrauch bei -20 °C gelagert. Die für die Untersuchungen verwendeten Fisch-DNA-Proben wurden freundlicherweise von J. MICHEL, Institut für Genetik der Justus-Liebig-Universität Gießen, zur Verfügung gestellt.[203]

Bei dem verwendeten Plasmid pXIR-2.5 handelt es sich um ein Derivat von pBluescript Sk(+) mit einer Länge von 5786 Bp mit einem inserierten 2.5 kBp langen *Xiphophorus*-Retrotransposon, das als XIR-2.5 bezeichnet wird. Das Plasmid wurde aus einer 500ml-Übernachtkultur eines transformierten *E. coli* Stamms HB 101 isoliert (siehe **2.4.1**). Die

Plasmid-DNA wurde durch einen quantitativen Restriktionsverdau mit *Eco*RV linearisiert. Die linearisierten Fragmente wurden mit dem QIAGEN PCR Purification Kit nach dem Protokoll des Herstellers aufgereinigt. Die Bakterienkultur wurde ebenfalls von J. MICHEL zur Verfügung gestellt.[203]

2.2.5 DNA-Proben von Saccharomyces cerevisiae

Die DNA-Proben von *S. cerevisiae* wurden aus 3ml-Kulturen des unten aufgeführten Stammes gewonnen, der freundlicherweise von J.H. HEGEMANN, Insitut für Mikrobiologie der Heinrich Heine Universität Düsseldorf, für diese Arbeit zur Verfügung gestellt wurde:

$$\#368\ \text{FY 1679, Genotyp a}/\alpha\quad \frac{\text{ura 3-52}}{\text{ura 3-52}}\ \frac{\text{GAL 2}}{\text{GAL 2}}\ \frac{\text{trp 1-63}}{+}\ \frac{\text{leu 2-}\Delta 1}{+}\ \frac{\text{his 3-}\Delta 200}{+}$$

Dazu wurden 3 ml autoklaviertes YPD-Medium mit einer Einzelkolonie angeimpft und über Nacht bei 28 °C im Schüttelwasserbad inkubiert, bis ein OD_{600}-Wert von 10 erreicht war. Die Isolation der DNA erfolgte wie in **2.4.1** beschrieben. Die isolierte DNA wurde bei -20 °C gelagert.

YPD-Medium:

20.0	g/l	Hefe-Extrakt
10.0	g/l	Pepton
20.0	g/l	Dextrose

2.3 Oligonukleotide

2.3.1 Planungskriterien für Oligonukleotide

Zur Planung von Oligodesoxyribonukleotiden oder kurz Oligonukleotiden, die als PCR-Primer dienen sollen, wurde das Programm Oligo 5.0 verwendet. Soweit keine speziellen Sequenzvorgaben oder Einschränkungen zu beachten sind, sollten die gewählten PCR-Primer dabei

- keine selbst- oder paarweise komplementären Sequenzen enthalten;
- weder im Plus- noch im Minus-Strang weitere exakte oder degenerierte Bindungsstellen in der Nähe der Ziel-Bindungsstelle im Templat besitzen (als Nähe ist der in der Sequenzdatenbank hinterlegte Sequenzkontext definiert, welcher die Zielsequenz enthält);
- keine oberhalb von 40 °C stabilen *stem-loop*- oder *hairpin*-Strukturen ausbilden;
- 3′-terminal die geringste interne Stabilität (ΔG) aufweisen;
- paarweise vergleichbare T_M-Werte (55±0.5 °C) aufweisen;

- ein PCR-Produkt einer Länge von 100-300 Bp amplifizieren.

Die Primer zur Amplifikation von cDNA zum Transkriptnachweis sollten paarweise mindestens in verschiedenen Exonen des betreffenden Gens lokalisiert sein, damit ein fälschlicherweise direkt aus genomischer DNA amplifiziertes Fragment anhand seiner größeren Länge identifiziert werden kann.

Die Planung von Oligonukleotiden, die als Hybridisierungssonden dienen sollen, unterliegt aufgrund des sehr eingeschränkten Sequenzraumes zwischen den Primerbindungsstellen starken Limitationen. Die Sonden können fälschlicherweise nicht als Primer dienen, da sie 3'-terminal entweder mit Fluoreszein markiert oder phosphoryliert sind. Für sie gelten generell ähnliche Planungskriterien wie für Primer; ihre T_M-Werte sollten jedoch diejenigen der Primer um ca. 5 K übersteigen.

2.3.2 Sequenzen und Bezeichnungen der verwendeten Primer

Nachfolgend sind alle in dieser Arbeit verwendeten Zielsequenzen und darauf hybridisierenden Primer in alphabetischer Reihenfolge aufgeführt. Die Primer wurden von der Firma MWG, Ebersberg, synthetisiert, die fluorophormarkierten Primer stammen von Applied Biosystems, Weiterstadt.

Tab. 2-9 **Zielsequenzen für die PCR-Amplifikationen.** Soweit nicht anders angegeben, handelt es sich um Zielsequenzen im humanen Genom. Die Genbank-Einträge bezeichnen die Zugriffsnummern zu den Datenbankeinträgen (accession number) der Sequenzen in der Gen-Datenbank „Nucleotide" des National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/).

Zielsequenz	Genbank	Bezeichnung / Genprodukt
BCR/ABL	AJ131466	Fusionstranskript der Gene für *breakpoint cluster region* / Abelsontyrosinkinase
D17S1818	AFM200zf4	Mikrosatellitenmarker
D17S800	AFMa312zc9	Mikrosatellitenmarker
F7c	J02933	Blutgerinnungsfaktor VIIc
HBB	AF007546	Betaglobin
HER-2/neu	M12036	Rezeptor für den humanen epidermalen Wachstumsfaktror 2
IGF-1	M12659	Rezeptor für den Insulinartigen Wachstumsfaktor 1
p16	U12820	Tumorsuppressorgen *p16* (*INK4A*)
p53	X54156	Tumorsuppressorgen *p53*
PAI-1	J03836	Plasminogen-Aktivator Inhibitor 1
rps3	U34347	Ribosomales Protein S3 aus *S. cerevisiae*
XIR	S77502	Retrotransposon XIR-2.5 aus *X. maculatus*
Xp53	U34751	Tumorsuppressorgen *p53* aus *X. maculatus*

Tab. 2-10 PCR-Primer. Bei Primern beruhen, soweit nicht explizit vermerkt, auf eigener Planung. Die für die **Primer** angegebenen Positionen beziehen sich auf die Datenbankeinträge (siehe Tab. 2-9). FAM: 5,6-Carboxyfluoreszein, HEX: 6-Carboxy-4,7,2′,4′,5′,7′-hexachlorofluoreszein, TET: 6-Carboxy-4,7,2′,7′-tetrachlorofluoreszein.

Zielsequenz	Name	Position	nt	Primersequenz 5′ → 3′
BCR/ABL	bcr(2)[1]	124	19	ACGTTCCTGATCTCCTCTG
	abl(1)[1]	380	20	AGATGCTACTGGCCGCTGAA
D17S1818	9a-TET[2]	17	23	**TET**-CATAGGTATGTTCAGAAATGTGA
	9m+[2]	138	25	GTTTCTTTGCCTACTGGAAACCAGA
D17S800	4a-HEX[2]	186	20	**HEX**-GGTCTCATCCATCAGGTTTT
	4m+[2]	33	30	GTTTCTTATAGACTGTGTACTGGGCATTGA
F7c	F7c-H1	77	20	GTCTGGAGGCTCTCTTCAAA
	F7c-H1-FAM	77	20	**FAM**-GTCTGGAGGCTCTCTTCAAA
	F7c-R1	195	20	AGAATTCCAAACCCCTAATG
	F7c-R2	203	20	AATTCCAAACCCCTAGGATA
	F7c-R3	230	20	GGCACCAACACTTCAAATAC
HBB	PC03[3]	143	20	ACACAACTGTGTTCACTAGC
	PC04[3]	233	20	CAACTTCATCCACGTTCACC
HER-2/neu	HER-H1	12	20	TGGAGGCCACAAGGTAAACA
	HER-H2	967	21	GAACTGGTGTATGCAGATTGC
	HER-R1	108	21	CCTACTCCATCCCAAGCCTAT
	HER-R2	1051	19	AGCAAGAGTCCCCATCCTA
IGF-1	IGF-H1	1135	16	AGCTCGGCATAGTCTT
	IGF-H2	1034	25	CCCAGGTACCCTTCTCCCAGAGTGG
	IGF-R1	1241	17	CCAAGTGAGGGGTGTGA
	IGF-R2	1139	21	GATGAGGCAAAGACTATGCCG
p16	p16-H	146	18	AAGCCATTGCGAGAACTT
	p16-R	253	18	CAGAGGGCAGAAAGAAAA
p53	p53-H	14674	19	TTCCTAGCACTGCCCAACA
	p53-R	14779	20	GACTGGAAACTTTCCACTTG
	QPCR-2a[2]	14513	21	CGGCGCACAGAGGAAGAGAAT
	QPCR-2b[2]	14834	21	CAAATGCCCCAATTGCAGGTA
	VNTR-1b[2]	8963	25	AAGAGCTGAGACTCCGTCTCAAAAT
	VNTR-2-HEX[2]	9053	24	**HEX**-ACAAAACATCCCCTACCAAACAGC
PAI-1	PAI-H1	116	18	AGTCTGGACACGTAGGGG
	PAI-H2	116	18	AGTCTGGACACGTAGGGA
	PAI-H3	38	18	CACCACCCCACCCAGCAC
	PAI-R1	232	18	CCAGCCACGTGATTGTCT
	PAI-R2-FAM	142	20	**FAM**-GGCCGCCTCCGATGATACAC
rps3	RPS3-H1	1376	20	CCAACCAAGACCGAAGTTAT

Zielsequenz	Name	Position	nt	Primersequenz 5′ → 3′
	RPS3-H2	1285	19	CGCTGACGGTGTCTTCTAC
	RPS3-R1	1531	16	GACAGCGGACAAACCA
	RPS3-R2	1646	20	CGGAAACAACAACTTCACAA
XlR-2.5	XlR-H[4]	1517	20	TGCTCGCTGCTGGCTGAACG
	XlR-R[4]	1739	20	TCGATGCGTCGAGGAGATAC
Xp53	xp53-H1	1010	19	CGGTGCTTCGAGGTCCGTG
	xp53-H2	1198	19	GAGCGCTCCTGCTCCAGAT
	xp53-R1	1377	19	CACAGATAACGATTACGGC
	xp53-R2	1267	19	GAGAGTGTAAATCTCCTTG

[1] A. FREIST[204]

[2] M. HAHN[205]

[3] C. WITTWER *et al.*[206]

[4] J. MICHEL[203]

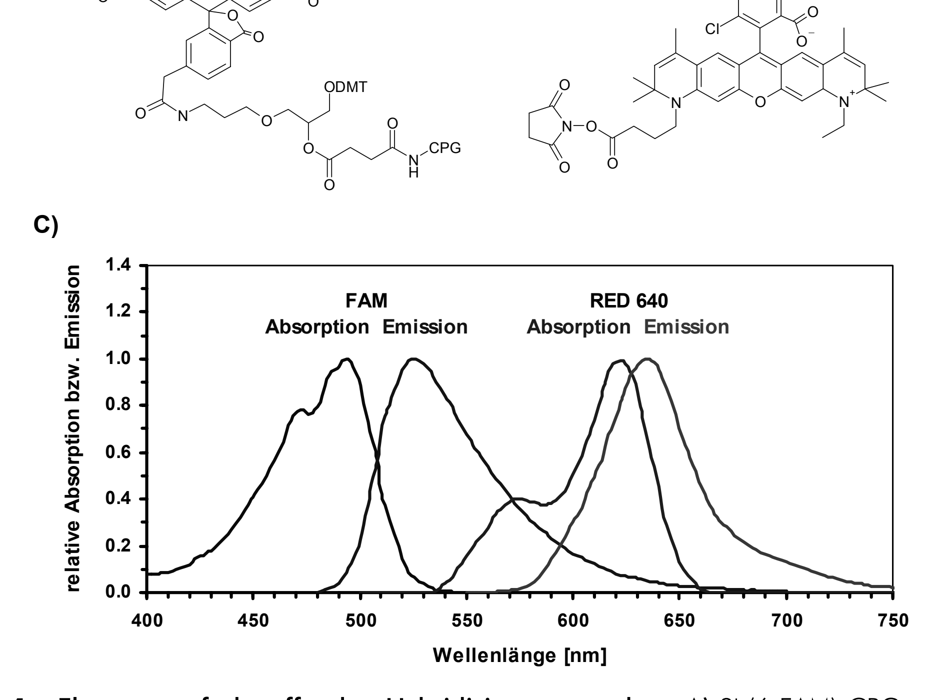


Abb. 2-1 Fluoreszenzfarbstoffe der Hybridisierungssonden. A) 3'-(6-FAM)-CPG zur 3′-terminalen Markierung, **B)** RED-640-N-Hydroxysuccinimidester zur 5′-terminalen Markierung von Oligonukleotiden. **C)** Absorptions-/Emissionsspektren der fluorophormarkierten Oligonukleotide. Die Spektren der 5 μM Lösungen der Hybridisierungssonden HBB-D (FAM) und HBB-A (RED-640) in Tris/HCl, pH 8.5 bei 20°C, wurden aufgezeichnet wie in 2.7 beschrieben.

2.3.3 Sequenzen und Bezeichnungen der verwendeten Hybridisierungssonden

In **Tab. 2-11** sind alle in dieser Arbeit verwendeten Paare von Hybridisierungssonden in alphabetischer Reihenfolge aufgeführt. Die Sonden wurden von der Firma TIB Molbiol, Berlin, synthetisiert. Die Donorsonden sind 3′-terminal mit Fluoreszein markiert, die Akzeptorsonden tragen 5′-terminal den Fluoreszenzfarbstoff LightCycler RED-640 und sind 3′-terminal phosphoryliert, damit sie von der DNA-Polymerase nicht verlängert werden können. Die Strukturformeln der zur Markierung verwendeten Fluoreszenzfarbstoffe sowie die Absorptions- und Emissionsspektren der an Oligonukleotide gekoppelten Fluorophore sind in **Abb. 2-1** gezeigt. Das Absorptionsmaximum von FAM liegt bei ca. 485 nm, das Emissionsmaximum bei 525 nm. RED-640 zeigt ein Absorptionsmaximum bei ca. 625 nm und ein Emissionsmaximum bei 635 nm. Das Emissionsspektrum des ebenfalls vom LightCycler (siehe **2.6.2**) detektierbaren Farbstoffs RED-705 ist in **Abb. 2-3** gezeigt.

Tab. 2-11 Hybridisierungssonden. Soweit nicht explizit vermerkt, beruhen die Sonden auf eigener Planung. Die für die Sonden angegebenen Positionen beziehen sich auf die Datenbankeinträge der Zielsequenzen (siehe Tab. 2-9). FAM: 5,6-Carboxyfluorescein, P: 3′-terminales Phosphat, RED: LightCycler RED-640.

Zielsequenz	Name	Pos.	nt	Sondensequenz 5′ → 3′
BCR/ABL	LC(1)[1]	154	35	GCAGAGTGGAGGGAGAACATCCGGGAGCAGCAGAA-**FAM**
	LC(2)[1]	190	26	**RED**-AAGTGTTTCAGAAGCTTCTCCCTGAC-**P**
HBB	HBB-D[2]	168	35	CAAACAGACACCATGGTGCACCTGACTCCTGAGGA-**FAM**
	HBB-A[2]	203	30	**RED**-GAAGTCTGCCGTTACTGCCCTGTGGGGCAA-**P**
HER-2/neu	HER-D1	36	34	ACATCCCCCTCCTTGACTATGCAATTTTACTAGA-**FAM**
	HER-A1	70	36	RED-GGATGTGGTGGGAAAACCATTATTTGATATTAAAAC-**P**
	HER-D2	992	29	GTATGCACCTGGGCTCTTTGCAGGTCTCT-**FAM**
	HER-A2	1021	27	**RED**-CCGGAGCAAACCCCTATGTCCACAAGG-**P**
IGF-1	IGF-D	1174	33	ATGAGACAGTGCCCTAAAGGGACCAATCCAATG-**FAM**
	IGF-A	1207	33	**RED**-CTGCCTGCCCCTCCATAGGTTCTAGGAAATGAG-**P**
p16	p16-D	185	30	CGGTAGGGACGGCAAGAGAGGAGGGCGGGA-**FAM**
	P16-A	215	34	**RED**-TGTGCCACACATCTTTGACCTCAGGTTTCTAACG-**P**
p53	P53-D	14707	34	CCCCAGCCAAAGAAGAAACCACTGGATGGAGAAT-**FAM**
	P53-A	14742	35	**RED**-TTTCACCCTTCAGGTACTAAGTCTTGGGACCTCTT-**P**

[1] A. FREIST[204]

[2] C. WITTWER *et al.*[206]

2.4 DNA-Präparationsverfahren

2.4.1 Präparation genomischer DNA aus biologischem Material

DNA aus biologischem Material (humane Blut- und Gewebeproben, *E. coli*, *S. cerevisiae*) wurde unter Verwendung von QIAGEN Kits (QIAGEN, Hilden) nach dem Protokoll des Herstellers durchgeführt. Das Standard-Protokoll wurde stets um einen Inkubationsschritt mit RNase A ergänzt.

Zur Isolation der DNA aus Blutproben wurde das QIAamp DNA Mini Kit „Blood“ verwendet, die DNA aus Gewebeproben wurde mit dem QIAamp DNA Mini Kit „Tissue“ isoliert. Die Plasmidextraktion erfolgte mit dem QIAGEN Plasmid Maxi Kit.

Die Hefe-Zellen wurden zunächst in Sorbitol-Puffer (1 M Sobitol, 100 mM EDTA, 14 mM 2-Mercaptoethanol) überführt und die Zellwände enzymatisch angedaut. Dazu wurden die 3ml-Kulturen zunächst 10 min bei 5000 g zentrifugiert. Das Pellet wurde anschließend in 600 µl Sorbitol-Puffer resuspendiert. Nach Zugabe von 200 U Zymolase wurde die Suspension 30 min bei 30 °C inkubiert, dann erneut 5 min bei 5000 g zentrifugiert und das Pellet in 180 µl ATL-Puffer des QIAamp Mini Kit „Tissue“ aufgenommen. Die weitere Prozedur der DNA-Isolation erfolgte entsprechend der Isolation aus Gewebeproben mit dem QIAamp DNA Mini Kit „Tissue“.

2.4.2 Isolierung von DNA-Fragmenten aus Polyacrylamidgelen

Zur Präparation von DNA-Fragmenten aus nicht-denaturierenden Polyacrylamidgelen (siehe **2.5.2**) wurden die Banden auf einem UV-Transilluminator (siehe **2.5.3**) mit einem sterilen Skalpell möglichst exakt aus dem Gel ausgeschnitten. Die DNA wurde dann über Nacht durch Inkubation der Gelstücke in dem 10- bis 20-fachen Volumen TE-Puffer aus der Gelmatrix herausgelöst. Die durch eine Sephadex-G-25-Gelfiltration mit einer NAP10-Säule von Pharmacia Biotech, Freiburg nach Herstellerangaben entsalzte Lösung wurde anschließend im Vakuum getrocknet und die DNA in Reinstwasser aufgenommen oder wie in **2.4.3** beschrieben aufgereinigt.

2.4.3 Isolierung von DNA-Fragmenten aus wässrigen Lösungen

Die Aufreinigung von PCR-Produkten aus den Reaktionsansätzen wurde unter Verwendung des QIAquick PCR-Purification-Kit (QIAGEN, Hilden) nach dem Protokoll des Herstellers durchgeführt.

2.4.4 Präparative DNA-Restriktionsspaltung

Zur präparativen Restriktionsspaltung wurden 1-50 μg DNA 3 h bei 37 °C mit 5 U Restriktionsenzym in einem Volumen von 50 μl des zugehörigen 1x Puffers des Herstellers inkubiert. Die Restriktionsfragmente wurden anschließend wie in **2.4.3** beschrieben aufgereinigt.

2.5 DNA-Analyse

2.5.1 UV/VIS-Spektrophotometrie

Die DNA-Konzentrationen von wässrigen Lösungen wurden UV-spektroskopisch bestimmt. Hierzu wurden Absorptionsspektren im Wellenlängenbereich von 220 - 320 nm gemessen (Hitachi U3000 UV-VIS Spektrophotometer mit Suprasil-Quarzglas Mikro-Küvetten (Helma, Müllheim/Baden)). Die Konzentrationsberechnung und Qualitätsabschätzung beruht auf den unten angegebenen Werten für Umrechnungsfaktoren und Stoffeigenschaften:[207]

- 1 $OD_{260} \cong 50\ \mu$g dsDNA bzw. 37 μg ssDNA;
- 660 g/mol als mittleres M_r für ein Basenpaar;
- das Verhältnis OD_{260}: OD_{280} einer reinen DNA-Lösung beträgt 1.8-1.9.

Ein OD_{260} ist definiert als die Menge eines Oligonukleotids, die in einem Volumen von 1 ml bei einer Schichtdicke von 1 cm und einer Wellenlänge von 260 nm eine Extinktion von 1 aufweist.

2.5.2 Polyacrylamidgel-Elektrophorese (PAGE)

15%iges Polyacrylamidgel:

14.5	% (w/v)	Acrylamid
0.5	% (w/v)	Bisacrylamid
0.08	% (w/v)	APS
0.2	% (v/v)	TEMED
in	(1x)-TPE-Puffer	

(5x)-AAP:

100	mM	EDTA, pH 8.0
25	% (w/v)	Ficoll 400
0.1	% (w/v)	Bromphenolblau
0.1	% (w/v)	Xylencyanol FF

(1x)-TPE-Puffer:

80	mM	Tris-Phosphat, pH 8.2
2	mM	EDTA

Zur Analyse von PCR-Produkten wurden 15%ige Polyacrylamidgele im Format 10 x 10 x 0.1 cm (Laufstrecke x Breite x Dicke) verwendet. Die Auftrennung erfolgte in einer Vertikalelektrophorese-Apparatur bei konstanter Stromstärke von 40 mA in einem TPE-Puffer-

system (Zusammensetzung der Gele und Puffer s. unten). Die Proben wurden vor dem Auftrag mit $^1/_5$ Endvolumen (5x)-AAP vermischt. Bei dem verwendeten Probenkamm mit 18 Taschen reichen 1-2 μl dieser Mischung aus, um nach EtBr-Färbung des Gels deutliche Produktbanden zu erhalten. Die EtBr-Färbung erfolgte in einer 1 μg/ml EtBr-Lösung, worin das Gel ca. 10 min auf einem Rotationsschüttler bewegt wurde. Anschließend erfolgte eine ebenfalls ca. 10 min Entfärbung in Reinstwasser.

2.5.3 Geldokumentation

EtBr-gefärbte DNA-Fragmente in Polyacrylamidgelen wurden auf einem Transilluminator unter Durchlicht bei 312 nm über ein Videodokumentationssystem dokumentiert. Die mit einer CCD-Kamera aufgezeichneten Graustufenbilder wurden im TIFF-Format gespeichert und können anschließend mit entsprechender Software weiterverarbeitet werden. Die Aufbereitung der Bilddateien bestand in der Feinjustierung von Helligkeit, Kontrast und Gamma sowie dem Zuschnitt oder der Vergrößerung der interessierenden Abschnitte. Die Nachbearbeitung der Gelbilder wurde mit dem Programm Photoshop durchgeführt.

2.5.4 Kapillarelektrophorese

Zur kapillarelektrophoretischen Analyse von Fragmentlängen und -mengen fluorophormarkierter Oligonukleotide wurde ein ABI Prism 310 Genetic Analyzer verwendet. Die Auftrennung der Fragmente erfolgte unter denaturierenden Bedingungen in einer mit flüssigem, linearem Polymer gefüllten Kapillare im elektrischen Feld nach elektrokinetischer Injektion. Das Gerät besitzt einen 10mW-Argonionenlaser mit den Emissionswellenlängen 488 nm und 514.5 nm zur Exzitation der die Anregungs-/ Detektionseinheit passierenden Fragmente. Die Fluoreszenzintensität wurde im Wellenlängenbereich zwischen 525 und 650 nm von einer CCD-Kamera aufgezeichnet und als Elektropherogramm mit der Auftragung der Fluoreszenzintensität gegen die Zeit (= Laufzeit der Fragmente bis zum Detektor) ausgegeben.

Jeweils 2 μl der mit Reinstwasser vorverdünnten Proben wurden zur Analyse zu 12 μl TSR gegeben, 3 min bei 95 °C im Heizblock denaturiert und dann auf Eis abgeschreckt. Die Vorverdünnung der Proben erfolgte so, dass die Fluoreszenzsignale im Messbereich der Kapillarelektrophorese-Apparatur liegen (zwischen 100 und 8000 relativen Fluoreszenz-Einheiten).

Die Proben wurden 5 s bei 15 kV elektrokinetisch injiziert und bei 15 kV und 60 °C in einer 47 cm x 50 μm (Länge x Durchmesser) Kapillare aus unbeschichteter gesinterter Kieselsäure aufgetrennt. Als Polymer wurde POP4 eingesetzt, das verwendete Laufmodul war „GS STR POP4 (1ml) C“. Die Laser-induzierten Fluoreszenz-Signale wurden unter Verwendung der ABI Prism Collection Software aufgezeichnet und anschließend mit der ABI Prism GeneScan Software ausgewertet.

Zur DNA-Fragmentlängen-Bestimmung wurden 0.5 μl GeneScan-500 TAMRA als interner

DNA-Fragmentlängenstandard zugegeben. Anhand der Fragmentlängen des Standards (siehe **Tab. 2-8**) wurde von der Software eine Kalibrierkurve nach der *„local Southern method"* erstellt: die Bestimmung der Proben-Fragmentlängen erfolgte hierbei anhand der Interpolation eines Polynoms zwischen den drei zum jeweiligen Probenfragment nächsten Standardfragmenten.

2.5.5 Schmelzkurvenanalyse

Zur Charakterisierung entstandener PCR-Produkte bei der RT-PCR mit SYBR-Green I-Detektion sowie zur quantitativen Analyse der Produkte bei der kompetitiven RT-PCR mit Detektion durch Hybridisierungssonden (siehe **3.4**) wurden Schmelzkurvenanalysen auf dem LightCycler (siehe **2.6.2**) durchgeführt.

Das Standard-Temperaturprofil zur Schmelzkurvenanalyse begann mit der Einstellung der Temperatur auf 72 °C (für SYBR-Green I) bzw. auf 55 °C (für Hybridisierungssonden). Anschließend wurde die Temperatur mit einem Gradienten von 0.2 K/s auf 95 °C erhöht, während die Fluoreszenzsignale der Proben im Modus *„continuous"* gemessen wurden. Dabei wurden die Proben so oft sequenziell gemessen, bis die Zieltemperatur erreicht war. Die Frequenz der Datenaufzeichnung je Probe hängt bei diesem Modus von der Anzahl der zu messenden Proben ab (die Messung einer Probe erfordert ca. 130 ms, siehe auch **2.6.2**).

Die aufgezeichneten Daten wurden anschließend entweder mit der LightCycler Data Analysis Software oder mit der selbst entwickelten Software (Algorithmus siehe **3.2**) ausgewertet. Die Temperaturwerte, bei welchen eine lokal maximale Signalabnahme auftritt, kennzeichnen die Schmelztemperatur (T_M) des Produkts bzw. des Produkt/Sonden-Hybrids.

2.6 Polymerase-Kettenreaktion (PCR)

2.6.1 Allgemeine Hinweise zur Durchführung

Bei der Durchführung von PCRs können Probleme auftreten, die durch entsprechende Maßnahmen vermieden werden können. Besonders zu erwähnen sind hier die Kontamination von Reaktionslösungen mit DNA-Templaten (z. B. PCR-Produkten, die bei der post-PCR-Analyse in hohen Konzentrationen vorliegen) sowie die Entstehung unerwünschter Nebenprodukte.

Die Wahrscheinlichkeit für auftretende Kontaminationen lässt sich minimieren, indem alle Reaktionskomponenten möglichst fern von PCR-Arbeitsplätzen aliquotiert, gelagert und pipettiert werden und die Arbeitsschritte vor und nach der PCR an räumlich getrennten Orten und mit separaten Geräten (z. B. Pipetten) durchgeführt werden.

Die Entstehung von unerwünschten Nebenprodukten bei der PCR ist nicht nur bedingt durch die Primersequenzen und -Konzentrationen sowie den Reaktionsbedingungen während der PCR (Hybridisierungstemperatur, Mg^{2+}-Konzentration, Polymerisationszeit; siehe **0**), sondern auch von den Bedingungen während des Ansetzens des Reaktionsgemisches. Hier ist darauf zu achten, dass alle Reagenzien eisgekühlt gemischt werden und als letzte Komponente das Enzym hinzugegeben wird. Nach der Fertigstellung des Reaktionsgemisches sollte die PCR unverzüglich gestartet werden. Eine weitere Erhöhung der Spezifität der Reaktion lässt sich durch die Anwendung einer *Hot-start*-Technik erreichen, bei welcher die Polymerase erst während der initialen Denaturierungsphase aktiviert oder zugegeben wird. Die in Teilen dieser Arbeit verwendete *Hot-start*-Technik beruht auf der Inaktivierung der Polymerase durch die Bindung eines Antikörpers. Während der initialen Denaturierung denaturiert auch der Antikörper und gibt die aktive Polymerase frei. In dieser Arbeit werden Primersysteme nur dann für Real-time-Quantifizierungen verwendet, wenn sie unter den Standardreaktionsbedingungen (**2.6.5**) in der PCR keine Nebenprodukte bilden. Die Systeme wurden während der Optimierungsphase im Anschluss an die PCRs anhand von PAGE-Analysen (**2.5.2**) und Schmelzkurvenanalysen (**2.5.5**) auf die Abwesenheit von Nebenprodukten überprüft.

Ein weiterer kritischer Punkt, insbesondere bei der quantitativen PCR mit externer Standardisierung, ist die Reproduzierbarkeit und Vergleichbarkeit von Reaktionen. Eine grundlegende Voraussetzung für eine gute Vergleichbarkeit von Reaktionen ist die möglichst identische Zusammensetzung der Reaktionsansätze. Um pipettierbedingte Variationen in den Konzentrationen der Reagenzien zwischen den zu vergleichenden Ansätzen zu minimieren, wurden bei allen Versuchen, wo immer möglich, die in allen Einzelreaktionen gleichen Bestandteile in einem *„master-mix"* gemeinsam pipettiert und anschließend in gleiche Volumina aufgeteilt.

2.6.2 Apparatetechnik

Zur konventionellen PCR wurde das GeneAmp 2400 PCR-System mit 0.2ml-Reaktionsgefäßen verwendet. Die Temperatursteuerung erfolgt bei diesem Gerät über Peltier-Elemente, die einen Aluminiumblock mit Vertiefungen für die Reaktionsgefäße temperieren. Eine Kondensation von Wasserdampf am Deckel der Reaktionsgefäße wird durch eine Deckelheizung verhindert.

Bei dem in dieser Arbeit verwendeten RT-PCR-Gerät handelt es sich um den ursprünglich von Idaho Technology (Idaho, USA) entwickelten LightCycler (siehe **Abb. 2-2**),[26,206,208] der 1999 von der Firma Boehringer Mannheim (jetzt Roche Diagnostics) übernommen und weiterentwickelt wurde. Da dieses Gerät verhältnismäßig neu ist und die methodischen Entwicklungen in dieser Arbeit speziell für dieses Gerät gemacht wurden, werden der Aufbau und die Funktion an dieser Stelle etwas ausführlicher erläutert.

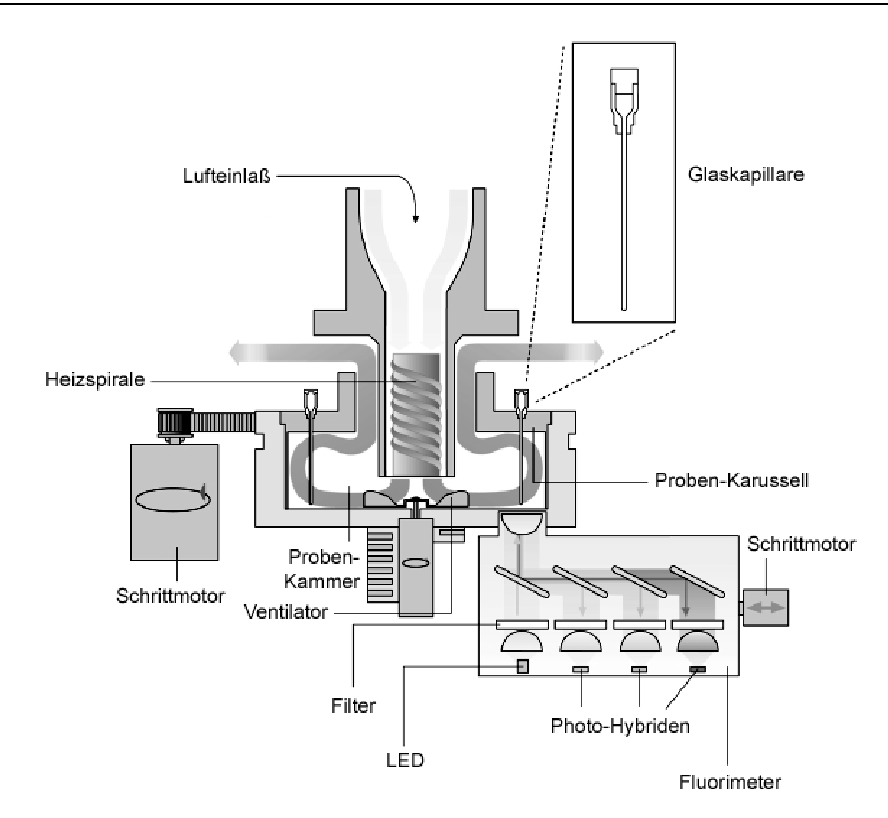


Abb. 2-2 **Schema des LightCyclers.** Die Symmetrieachse der Probenkammer befindet sich in der Schnittebene. Einzelheiten zum Aufbau und die Erklärung der Funktionsweise sind im Text gegeben. Abbildung verändert gegenüber dem Original aus dem LightCycler-Benutzerhandbuch.

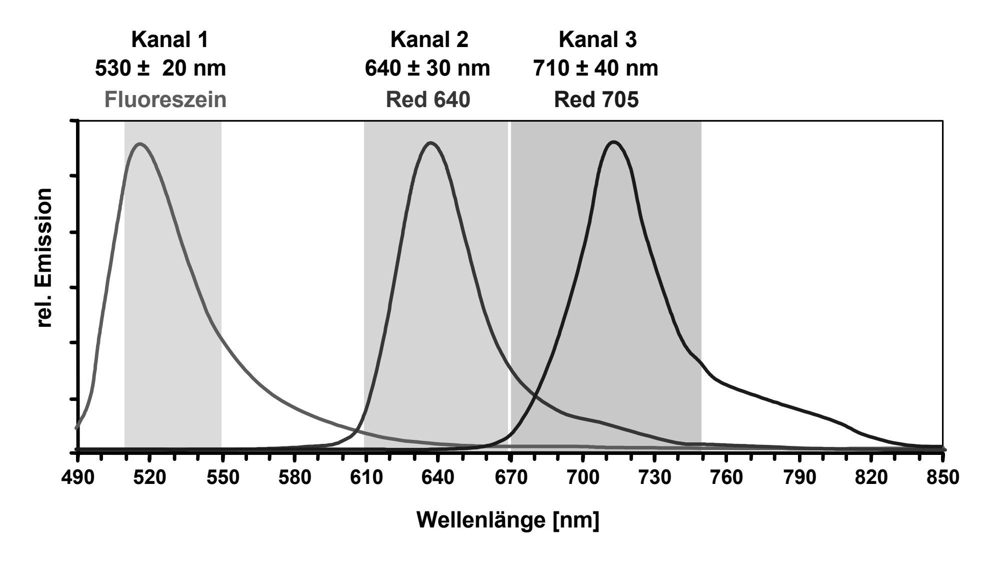


Abb. 2-3 **Emissionsspektren der im LightCycler verwendeten Fluoreszenzfarbstoffe und Detektionsbereiche der Kanäle.** Die Detektionsbereiche sind farbig unterlegt. Sie entsprechen der doppelten Halbbandbreite um den Bandpass der Interferenzfilter. Verändert aus dem LightCycler-Benutzerhandbuch.

Das Gerät ist im Wesentlichen eine Kombination aus einer temperierbaren Probenkammer und einem Fluorimeter. Die Temperierung erfolgt durch einen Luftstrom der Umgebungsluft, welche durch eine Heizwendel von Raumtemperatur bis auf über 100 °C aufgeheizt werden kann. Der temperierte Luftstrom wird durch eine rotationssymmetrische Probenkammer geleitet, in welcher ein Ventilator für zusätzliche Verwirbelungen und somit für eine gleichmäßigere Temperaturverteilung sorgt.

Die Reaktionsansätze mit einem Maximalvolumen von 20 µl befinden sich in Glaskapillaren, welche in einem Proben-Karussell fixiert werden, das in die temperierte Probenkammer eingesetzt wird. Das Proben-Karussell bietet Platz für bis zu 32 Glaskapillaren. Durch die Kombination aus Luft als Temperierungsmedium mit geringer spezifischer Wärmekapazität einerseits und einem zylindrischen Probenvolumen mit hohem Verhältnis von Oberfläche zu Volumen andererseits sind Heiz- und Kühlraten der Proben von bis zu 20 K/s möglich, wobei die Temperaturverteilung im gesamten Probenvolumen jederzeit sehr gleichmäßig ist. Hierdurch können sehr kurze Zykluszeiten für die Amplifikation erreicht werden. Während eine PCR mit 35 Amplifikationszyklen in herkömmlichen Geräten mit Peltier-Element bestückten Aluminiumblöcken („Block-*Cycler*") üblicherweise mehrere Stunden dauert, ist sie in einem LightCycler schon nach 15-30 min abgeschlossen. Dieses auch als *rapid-cycle*-PCR bezeichnete Verfahren eignet sich sehr gut zur schnellen Optimierung und sukzessiven Analyse verschiedener Reaktionen mit unterschiedlichen Temperaturprotokollen und kompensiert den im Vergleich zu einem „Block-*Cycler*" mit 96 Probenplätzen geringeren Probendurchsatz pro Lauf für Hochdurchsatz-Bestimmungen.

An einer Stelle im Boden der Probenkammer befindet sich das Linsensystem der Optik zur Anregung und Detektion der Fluoreszenz in den Reaktionsansätzen. Das Anregungslicht mit einer Wellenlänge von 470 nm wird von einer LED erzeugt und über das Linsensystem auf die Kapillarenspitze fokussiert. Das im Inneren der Kapillare daraufhin von den im Reaktionsgemisch enthaltenen Fluorophoren emittierte Fluoreszenzlicht wird zum größten Teil durch Totalreflexion an den Glaswänden zur Kapillarenspitze zurückgeleitet und von ihr wiederum auf das darunter befindliche Linsensystem fokussiert. Über dichroitische Filter wird es dann an Photohybrid-Detektoren weitergeleitet. Die optische Einheit enthält drei Interferenzfilter für Licht der Wellenlängen von 530, 640 und 710 nm, so dass eine simultane Detektion von drei verschiedenen Fluoreszenz-Farbstoffen mit entsprechenden Emissionsmaxima möglich ist. Eine Übersicht über die Emissionsspektren der im LightCycler verwendeten Fluoreszenzfarbstoffe sowie den durch die Filter bedingten Detektionsbereichem (=Kanäle) ist in **Abb. 2-3** gegeben.

Da sich immer nur eine Glaskapillare über dem Linsensystem befinden kann, können die Proben nur sequenziell gemessen werden. Zur Messung einer Probe wird das Proben-Karussell mit Hilfe eines Schrittmotors so gedreht, dass die Glaskapillare mit der zu messenden Probe zentrisch über dem Linsensystem positioniert wird. Die Positionierung dauert ca. 130 ms, der Messvorgang 20 ms, so dass zur sequenziellen Messung eines mit 32 Proben vollbesetzten Probenkarussells etwa 5 s benötigt werden. Die Fluoreszenz kann entweder einmal pro Zyklus zu unterschiedlichen Zeitpunkten gemessen werden oder in Form eines kontinuierlichen Messvorgangs, z. B. zum Aufnehmen einer Schmelzkurve.

2.6.3 Reaktionskomponenten

Alle Reaktionskomponenten wurden in Aliquots für 10-100 Reaktionen bei -20 °C gelagert und zum Gebrauch auf Eis aufgetaut und pipettiert.

Bei der verwendeten thermostabilen Polymerase handelte es sich um die rekombinant in *E. coli* produzierte *Taq* DNA-Polymerase aus *Thermus aquaticus* oder das Stoffel-Fragment, welches der *Taq* DNA-Polymerase ohne die endonukleolytische Domäne entspricht. Die Enzyme lagen in einer Konzentration von 5 U/µl in 50 mM Tris/HCl, pH 7.5, 0.1 mM EDTA, 5 mM DTT und 50 % Glycerin vor. Es wurde der vom Hersteller zusammen mit der Polymerase gelieferte PCR-Puffer verwendet. Der (1x)-PCR-Puffer enthielt 10 mM Tris/HCl, pH 8.3, 50 mM KCl sowie 1.5 mM $MgCl_2$. Die Magnesium-Konzentration im Reaktionsansatz wurde durch Zugabe von $MgCl_2$-Lösung so eingestellt, dass eine optimale Amplifikationseffizienz bzw. ein optimaler Signalverlauf bei minimaler Nebenproduktbildung erreicht wurde (siehe **0**).

Als 2'-Desoxynukleosid-5'-triphosphat-Gemisch (dNTP-Mix) wurde eine äquimolare Stammlösung von je 20 mM dATP, dCTP, dGTP und dTTP eingesetzt. Der dNTP-Mix wurde aus 100 mM Lösungen der Desoxyribonukleotide hergestellt.

Die in dieser Arbeit verwendeten und in **2.3.2** aufgeführten Primer wurden mit Reinstwasser auf eine Konzentration von 5 µM eingestellt und aliquotiert. Die Hybridisierungssonden (siehe **0**) wurden in Aliquots mit Konzentrationen von 2 µM und 4 µM in TE-Puffer, pH 8.0, vom Hersteller bezogen.

Die SYBR-Green I-Stammlösung wurde 1:100 mit Reinstwasser verdünnt und aliquotiert gelagert. Von diesen Verdünnungen wurden Arbeitslösungen mit einer Endverdünnung von 1:1000 angesetzt, die direkt in die PCR-Reaktionen eingesetzt wurden.

2.6.4 Standardreaktionsbedingungen der konventionellen PCR

Soweit nicht anders angegeben, wurden die konventionellen PCRs unter den unten aufgeführten Bedingungen durchgeführt: Das Ansatzvolumen betrugt immer 25 µl. Für präparative Produkt-Generierung wurden entsprechend größere Reaktionsansätze hergestellt, zur Amplifikation allerdings in Aliquots zu je 25 µl aufgeteilt. Das Standard-Temperaturprogramm einer konventionellen PCR ist in **Tab. 2-12** wiedergegeben.

Tab. 2-12 Standard-Temperaturprogramm einer konventionellen PCR

Programm	Anzahl Wdh.	Segment Nr.	Temperatur [°C]	Dauer [s]
1 Initiale Denaturierung	1 x	1	95	300
2 Amplifikation	35 x	1	95	30
		2	55	30
		3	72	60
3 Finale Extension	1 x	1	72	120

Reaktionsansatz:

20 - 50	ng	Templat-DNA
0.5	µM	Hin-Primer
0.5	µM	Rück-Primer
200.0	µM	dNTP-Mix
0.5	U	DNA-Polymerase

in (1x)-PCR-Puffer (enthält 1.5 mM $MgCl_2$)

2.6.5 Standardreaktionsbedingungen der Real-time-PCR

Soweit nicht anders angegeben, wurden die RT-PCRs unter den im unten aufgeführten Bedingungen durchgeführt.

Reaktionsansatz mit SYBR-Green I:

0 - 30	ng	Templat-DNA
6.0	mM	$MgCl_2$
0.5	g/l	BSA
0.5	µM	Hin-Primer
0.5	µM	Rück-Primer
200.0	µM	dNTP-Mix
1:30000	x	SYBR-Green I
0.5	U	DNA-Polymerase

in (1x)-PCR-Puffer

Reaktionsansatz mit Hybridisierungssonden:

0 - 30	ng	Templat-DNA
6.0	mM	$MgCl_2$
0.5	g/l	BSA
0.5	µM	Hin-Primer
0.5	µM	Rück-Primer
0.2	µM	Donorsonde
0.4	µM	Akzeptorsonde
200.0	µM	dNTP-Mix
0.5	U	DNA-Polymerase

in (1x)-PCR-Puffer

Das Ansatzvolumen betrug 10 µl.

Das Standard-Temperaturprogramm einer RT-PCR ist in **Tab. 2-13** wiedergegeben. Der eingestellte Temperaturgradient betrug immer 20 K/s. Der Analyse-Modus des 1. Programms war „*None*", der des 2. Programms war „*Quantification*". Der Zeitpunkt der Fluoreszenzmessung war für SYBR-Green I-Detektion am Ende des Segments 3 von Programm 2 (nach der Polymerisation), für die Detektion mit Hybridisierungssonden am Ende des Segments 2 von Programm 2 (nach der Hybridisierung). Der eingestellte Messmodus war „*Single*", d. h. in jedem Zyklus wurde jeweils ein Messpunkt aufgezeichnet.

Tab. 2-13 Standard-Temperaturprogramm einer RT-PCR.

Programm	Anzahl Wdh.	Segment Nr.	Temperatur [°C]	Dauer [s]
1 Initiale Denaturierung	1 x	1	95	30
2 Amplifikation	50 x	1	95	0
		2	55	5
		3	72	10

Dem Temperaturprogramm in **Tab. 2-13** konnte sich für die Schmelzkurvenanalyse nach der Amplifikation ein dritter Programmpunkt „Schmelzkurve“ anschließen. Das Temperaturprofil dafür ist in **2.5.5** beschrieben. Der Analyse-Modus dieses Programms war „*Melting Curves*“.

2.6.6 *Hot-start*-PCR

Zur Erhöhung der Spezifität der Produktbildung während der PCR wurde in Teilen dieser Arbeit eine *Hot-start*-Technik mit Antikörpern gegen die *Taq* DNA-Polymerase verwendet. Dazu wurde die DNA-Polymerase vor der Herstellung des PCR-Reaktionsgemisches über einen Zeitraum von 5 min bei RT mit dem 27-fachen molaren Überschuss an Antikörper im Antikörper-Verdünnungspuffer des Herstellers (TaqStart-Dilution Buffer) inkubiert. Die Enzymkonzentration in diesem Gemisch betrug 0.5 U/µl. Dem PCR-Reaktionsansatz wurde sodann die mit dem Antikörper präinkubierte *Taq* DNA-Polymerase zugesetzt.

2.6.7 Optimierungsstrategie

Die Parameter der PCRs wurden so optimiert, dass bei maximaler Spezifität eine möglichst hohe Effizienz der Amplifikation bzw. Signalbildung erreicht wurde. Der erste Schritt dieser Strategie bestand in der Wahl geeigneter Primerpaare. Die Sequenzen der Primer wurde mit Hilfe spezieller Software so gewählt, dass neben besonderen Sequenzvorgaben die in **2.3.1** genannten Planungskriterien erfüllt waren. Je Zielsequenz wurden wenn möglich mehrere miteinander kompatible Hin- und Rück-Primer gewählt und in allen Kombinationsmöglichkeiten unter Standard-Reaktionsbedingungen für eine RT-PCR eingesetzt. Das Primerpaar, welches bei maximaler Spezifität die größten Produktmengen liefert bzw. die größte Amplifikationseffizienz aufweist, wurde weiterverwendet. Der nächste Schritt bestand in der Optimierung der Mg^{2+}-Konzentration. Hier wurden die Effizienz und Spezifität der Produktbildung des gewählten Primerpaares unter verschiedenen Magnesium-Ionenkonzentrationen verglichen. Schließlich wurde im letzten Schritt, sofern nötig, die Temperatur für die Hybridisierung der Primer justiert. Bei den im LightCycler durchgeführten *rapid-cycle*-PCRs war es darüber hinaus möglich, die Bildung zu langer Produkte zu verhindern, indem die Polymerisierungszeiten verkürzt wurden.

Enzym, Primer, dNTPs und Sonden wurden, soweit nicht anders angegeben, im Überschuss eingesetzt. Ihre Konzentrationen wurden nicht optimiert. Eine Verringerung der Konzentrationen, insbesondere von Enzym und Primer, führte schnell zu einer deutlichen Verschlechterung der Effizienz, wohingegen eine (mäßige) Konzentrationserhöhung keinen signifikanten Einfluss auf die Reaktion zeigte.

2.7 Fluoreszenzspektroskopie

Für die Aufzeichnung von Emissions- und Exzitationsspektren von Fluoreszenzfarbstoffen und fluorophormarkierten Oligonukleotiden wurde das F-4500 Fluoreszenz-Spektrophotometer mit Suprasil-Quarzglas Ultra-Mikro-Fluoreszenzküvetten (Helma, Müllheim/Baden) verwendet. Von den Rohdaten wurde mit Excel der gleitende Mittelwert zur Glättung berechnet.

2.8 Programmierung

Die Entwicklung der Software erfolgte auf einem PC mit 1 GHz Pentium-Prozessor und 256 MB Arbeitsspeicher unter Windows 98 als Betriebssystem. Als Entwicklungsumgebung wurde Borland Delphi, Version 6 (Standard Edition) verwendet.

Die Programmierung bestand in der objektorientierten Formulierung des Problems, der Konzeption adäquater Objekte zur Datenverarbeitung und der visuellen Bediener-Schnittstelle. Diese umfasste die Anwendungsfenster oder Formulare sowie die darin befindlichen Ein- und Ausgabe-Komponenten. Die erstellten Objekte zur Datenverarbeitung beinhalten Methoden zum Laden, Berechnen und Analysieren der Daten. Die zu ladenden Daten befanden sich in den von der LightCycler Run Software generierten Messwert-Dateien in Form von tabellarisch organisierten Textdaten im ASCII-Format (*American standard code for information interchange*). Pro Experiment werden von der LightCycler Run Software jeweils drei Dateien generiert: (i) eine Informationsdatei mit Angaben über Protokoll und Proben (*about file* „ABT“), (ii) eine Messwertdatei mit Aufzeichnungen von Zeit, Zyklusnummer, Temperatur und Fluoreszenzwerten (*fluorescence-file*, „FLO“) sowie (iii) einer Referenzdatei mit der Aufzeichnung einer kontinuierlichen Temperaturmessung während des gesamten Experiments (*temperature file*, “TEM“). Zur Auswertung des Experiments wurden nur die ABT- und die FLO-Dateien benötigt und geladen.

Die Gestaltung der Bedienerführung richtete sich nach der von IBM festgelegten CUA-Norm (*common user access*), die Teil der SAA (*system application architecture*) ist. Dies betraf insbesondere die Menüstruktur, Tastaturbelegung, Dialogboxgestaltung sowie die Funktionalität der Maus.

2.9 Statistische Tests

Zur Prüfung, ob zwei Stichproben Teil derselben Grundgesamtheit sind, wurde für normalverteilte Werte der t-Test angewendet mit einer Irrtumswahrscheinlichkeit von 0.05. Sind die Werte der Stichproben jedoch nicht normalverteilt, liefert der t-Test keine sinnvollen Ergebnisse. Wenn die Verteilung der Werte innerhalb der Stichproben nicht bekannt ist, müs-

sen sogenannte nicht-parametrische Tests angewendet werden, die keine Annahmen über die Art der zu Grunde liegenden Verteilungen machen. War die Verteilungsfunktion der Werte unbekannt bzw. nicht-normal, wurde der *pair-wise randomization test* verwendet.[209] Dieser Test ist ein sehr einfaches nichtparametrisches Verfahren, welches keinerlei Annahmen über die zu Grunde liegende Verteilungsfunktion macht.

Beim *pair-wise randomization test* wird die Wahrscheinlichkeit berechnet, mit welcher der gemessene Unterschied der Mittelwerte eines Stichprobenpaares nur zufällig ist. Sei A eine Stichprobe vom Umfang n_A, B eine Stichprobe vom Umfang n_B und $\left|\overline{A}-\overline{B}\right|$ der beobachtete Unterschied der Mittelwerte. Die insgesamt n_A+n_B Messwerte beider Stichproben werden zum Test als Grundgesamtheit aufgefasst. Daraus werden alle Kombinationen $\left(n=\frac{(n_A+n_B)!}{n_A!\cdot n_B!}\right)$ von Stichproben A' und B' des gleichen Umfangs wie die ursprünglichen Stichproben gebildet. Der Anteil der durch Neukombination gewonnenen Stichproben, deren Unterschied in den Mittelwerten mindestens so groß ist wie der Unterschied der Mittelwerte der ursprünglichen Stichproben, p, ist ein Maß für die Wahrscheinlichkeit, mit welcher der beobachtete Unterschied in den Mittelwerten nur zufällig ist. Unterschiede wurden als signifikant beurteilt, wenn $p \leq 0.05$ war. Der Rechenaufwand für den Test steigt exponentiell mit dem Stichprobenumfang. Der Test ist für Stichprobenumfänge von 10 bis 15 Werten noch gut anwendbar (mit $n = 155\cdot10^6$ Permutationen für $n_A = n_B = 15$). Bei größeren Stichproben kann man einen repräsentativen Teil aller möglichen Permutationen prüfen.

Als einfacher Test zur Feststellung von Normalverteilung wurde folgendes Verfahren benutzt: Die normierten Daten (= Anzahl Beobachtungen je Klasse geteilt durch die Gesamtzahl der Beobachtungen) werden in Form eines kumulativen Histogramms dargestellt. Von diesen Daten wurde die Quantile der Standard-Normalverteilung berechnet (die Quantile entspricht dem X-Wert, bei welchem die kumulierte Verteilungsfunktion den gegebenen Funktionswert aufweist. Bei der Normalverteilung ist die Quantile von 0: $-\infty$, von 0.5: ± 0 und von 1: $+\infty$). Die Quantilen wurde dann gegen die Werte der kumulierten Standard-Normalverteilung der Klassengrenzen abgetragen. Dies entspricht im Prinzip der Abtragung der kumulierten Häufigkeiten in einem Diagramm, dessen Werteachse die Skalierung einer kumulativen Normalverteilung aufweist. Dadurch ergibt sich für den Fall, dass die Werte normalverteilt sind, eine Gerade. Ebenso wurden z. B. exponentielle Funktionen in Diagrammen mit logarithmischer Werteachse als Geraden abgebildet. Der Korrelationskoeffizient der Regressionsgeraden durch diese Punkte gibt Aufschluss über die Wahrscheinlichkeit, mit welcher die Werte normalverteilt sind.

3 ERGEBNISSE SOWIE DEREN DISKUSSION

3.1 Charakterisierung von Real-time-PCR-Systemen

Die Grade des Falschen festzustellen und die Nothwendigkeit des Grundirrthums als der Lebensbedingungen des vorstellenden Seins – Aufgabe der Wissenschaft.

Friedrich Nietzsche

Die RT-PCR ermöglicht eine vergleichsweise einfach durchzuführende Quantifizierung von Nukleinsäuresequenzen innerhalb großer Konzentrationsbereiche. Die experimentell bestimmte Größe, der C_T-Wert, ist dabei proportional zum Logarithmus der Templatmenge im Reaktionsansatz. Bei der Berechnung der Ergebnisse steht die fehlerbehaftete Größe somit im Exponenten, was die erreichbare Präzision empfindlich beeinflusst. Es ist daher bei RT-PCR-Experimenten nicht nur von Bedeutung, dass nur das spezifische Produkt entsteht, sondern die Amplifikation auch mit reproduzierbar hoher Effizienz erfolgt und die Signalkurven ein hohes Signal-Rausch-Verhältnis aufweisen. In den folgenden Abschnitten wird untersucht, welche Parameter die Qualität der Signalkurven beeinflussen.

3.1.1 Signalverlauf während der PCR

Die aufgezeichneten Signalkurven lassen sich augenscheinlich in 4 bis 5 Phasen unterteilen (**Abb. 3-1**). Während der ersten Zyklen ist das Signal deutlich größer als Null, nimmt aber

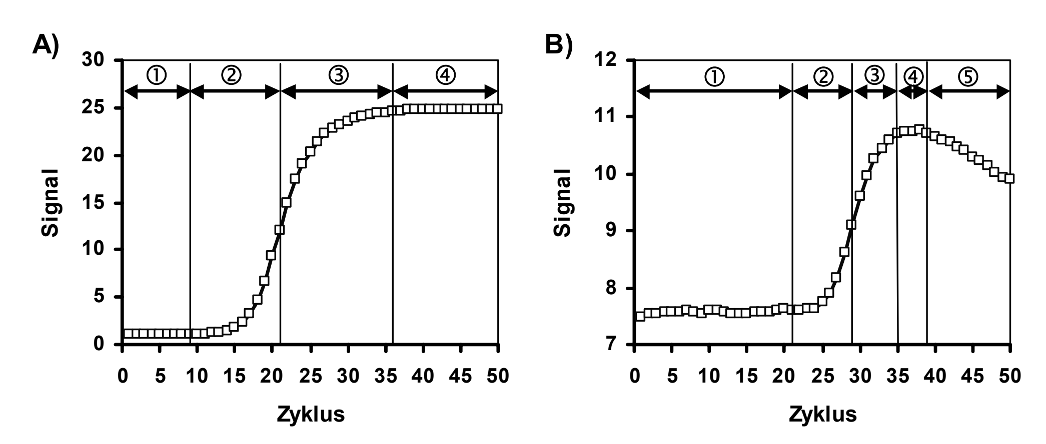


Abb. 3-1 Signalverlauf während einer PCR. A) SYBR-Green 1-Detektion. Die Zielsequenz *RPS3* wurde aus 105 initialen Kopien genomischer DNA von *S. saccharomyces* amplifiziert. **B)** Hybridisierungssonden-Detektion. Amplifiziert wurde die *HBB*-Zielsequenz ausgehend von 104 initialen Kopien humaner genomischer DNA. Abschnitte der Kurve: ① Basislinie, ② exponentielle Phase, ③ Übergang in die Plateauphase, ④ Plateauphase, ⑤ Depressionsphase.

nicht exponentiell zu. Es folgt die Phase erkennbarer exponentieller Signalzunahme, eine Phase abnehmender Effizienz als Übergang in die Plateauphase und schließlich die Plateauphase.

Bei Messungen mit sequenzspezifischen Sonden, nicht aber mit SYBR-Green 1, folgt manchmal noch eine Depressionsphase, in welcher das Signal wieder abnimmt und sich einem konstanten Grenzwert annähert. Dieses Verhalten ist vermutlich bedingt durch die relativ zur entstandenen Produktmenge geringen Sondenmenge im Ansatz. Durch die Massenwirkung rehybridisieren die aufgeschmolzenen Produktstränge schneller mit ihren Komplementen als mit den Sondenpaaren. Bis die Produktneubildung in den folgenden Zyklen schließlich auf Null zurückgeht, fällt die Signalintensität ab. Das Ausmaß des späten Signalabfalls ist abhängig von der Zielsequenz, den Sondensequenzen sowie den Versuchsbedingungen.

Das während der PCR aufgezeichnete Fluoreszenzsignal bzw. die daraus abgeleiteten Signalgrößen sind also nicht immer proportional zur Menge an PCR-Produkt. Bis einschließlich der exponentiellen Phase ist diese Beziehung aber gültig. Die Signalhöhe in den der exponentiellen Phase folgenden Zyklen kann allerdings nur sehr bedingt dazu verwendet werden, die Menge entstandener PCR-Produkte abzuschätzen. Da die proportionale Beziehung zwischen der Menge an PCR-Produkt und Signalintensität während der exponentiellen Phase aber gegeben ist, hat die Produktrehybridisierung in den späteren Zyklen der PCR keinen Einfluss auf die Quantifizierung der initialen Templatmengen. Die Magnesium-Ionenkonzentrationen des Standard-Reaktionsansatzes für die in dieser Arbeit durchgeführten RT-PCRs betragen 6 mM. Dies ist eine deutlich höhere Konzentration, als sie üblicherweise in konventionellen PCRs eingesetzt wird. Zwar kann eine hohe Magnesium-Ionenkonzentration durch eine Erhöhung der Schmelztemperaturen die Spezifität der Primerbindung erniedrigen, doch zeigt sich, dass sie einen positiven Einfluss auf die Signalkurven hat (siehe **Abb. 3-2**).

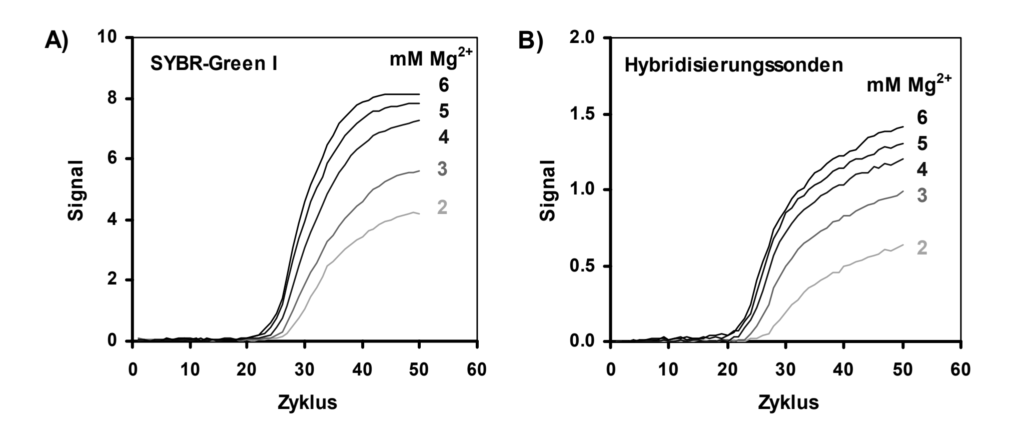


Abb. 3-2 **Signalverlauf während der PCR in Abhängigkeit von der Magnesium-Ionenkonzentration.** Die Beispiele zeigen die Amplifikationskurven von p53-H/R bei unterschiedlichen Magnesium-Konzentrationen. Als Templat wurden 30 ng humane genomische DNA eingesetzt. **A)** SYBR-Green I Detektion. **B)** Detektion mit Hybridisierungssonden.

Unabhängig vom Detektionsformat können in Abhängigkeit von der Mg^{2+}-Konzentration insgesamt höhere Signalintensitäten erreicht werden, wobei die exponentiellen Phasen mit zunehmender Mg^{2+}-Konzentration nach links verschoben sind, was auf eine höhere Amplifikationseffizienz hinweist. Durch die höheren Signalintensitäten vergrößern sich das Signal/Rausch-Verhältnis sowie der Bereich der erkennbaren exponentiellen Phase. Aus diesen Gründen sind bei höheren Magnesium-Ionenkonzentrationen exaktere Quantifizierungen möglich.

3.1.2 Einfluss fehlerhafter C_T-Bestimmungen auf die Quantifizierung

Die Datenpunkte der exponentiellen Phase der Amplifikationskurven lassen sich verwenden, um die C_T-Werte zu berechnen. Dazu müssen die in diskreten Zeitintervallen aufgezeichneten Daten interpoliert werden, da der Schwellenwert in der Regel nicht exakt einer gemessenen Signalhöhe entspricht und die Amplifikationskurven somit zwischen den Datenpunkten schneidet. Abweichungen der Signalanteile von dem zur Produktmenge proportionalen Wert beeinflussen diese Interpolation und damit die Lage der C_T-Werte. Da die C_T-Werte proportional zum Logarithmus der initialen Templatmenge sind (siehe **Gl. 1.7**), gehen Fehler in der Bestimmung der C_T-Werte exponentiell in die Berechnung der initialen Templatmenge ein.

Die Größenordnung des Einflusses in Abhängigkeit verschiedener Parameter ist anhand von Simulationen zur Berechnung von C_T-Werten und initialen Templatmengen gezeigt (**Abb. 3-3**). In diesen Simulationen wurden exponentielle Signalzunahmen über drei bis sechs Zyklen gemäß der angegebenen initialen Templatmenge und Amplifikationseffizienz berechnet, wobei dem berechneten Signalwert ein normalverteilter Fehler mit der angegebenen Streuung addiert wurde. Der Fehler ist hierbei unabhängig von der Signalhöhe. Die C_T-Werte wurden als Schnittpunkte der Regressionsgeraden durch die berechneten Punkte mit dem angegebenen Schwellenwert berechnet. Die berechneten C_T-Werte sind normalverteilt, so dass als Maß für den Fehler die Standardabweichung (SD) der ermittelten C_T-Werte und initialen Templatmengen herangezogen wurde.

Die Streuung der C_T-Werte (SD_{C_T}) nimmt linear mit der Signalfehlerstreuung zu (**Abb. 3-3 A**). Bei einer für den LightCycler durchschnittlichen Signalfehlerstreuung von 20 % des Schwellenwertes (MW aus mehr als 100 Experimenten mit SYBR Green I und Hybridisierungssonden; in der Simulation entspricht dies einem Parameterwert von $f = 200$ für die Fehlerfunktion), der sich am unteren Ende der exponentiellen Phase befindet, ist $SD_{C_T} = 0.142$ und die berechneten Templatmengen schwanken zwischen 70 % und 144 % des wahren Wertes. Die Verteilung der berechneten Templatmengen hat einen Variationskoeffizient CV von ca. 10 % (siehe rote Markierung in **Abb. 3-3**). Höhere Schwellenwerte verringern die Streuung der C_T-Werte (**Abb. 3-3 B**), sofern die Signalfehlerstreuung unabhängig oder zumindest nicht exponentiell von der absoluten Signalintensität ab-

hängig ist. Ebenso verringern auch höhere Effizienzen die Werte für SD_{C_T} (**Abb. 3-3 C**), wobei allerdings hier, bedingt durch den exponentiellen Zusammenhang von ε und N_0 ($N_0 = S \cdot \varepsilon^{-C_T}$), CV mit steigender Effizienz trotzdem leicht zunimmt. Während CV bei einer Effizienz von 1.7 bei ca. 10 % liegt, steigt er bei einer Effizienz von 2.0 auf über 11.6 % an. Die Verwendung einer zunehmenden Anzahl von Datenpunkten für die Regressionsanalyse (n) bewirkt eine bessere Glättung bzw. Filterung der Signalfehlerstreuung und dadurch auch eine Verringerung von SD_{C_T} (**Abb. 3-3 D**).

Aus den Ergebnissen der Simulationen folgt, dass ein alleine durch das Signalrauschen bedingter relativer Standardfehler in der Bestimmung der initialen Templatmengen bei der Schwellenwertmethode von etwa 10 % für das LightCycler-System nicht unterschritten werden kann. Abgesehen vom gerätebedingten Signalrauschen, worauf der Experimentator in der Regel keinen Einfluss hat, ist die Lage des Schwellenwertes der Parameter mit dem größten Einfluss auf den Standardfehler. Je höher der Schwellenwert gewählt wird, desto geringer ist der Einfluss des statistischen Signalrauschens auf das Ergebnis. Da die Stärke des Signalrauschens tatsächlich jedoch leicht mit der absoluten Signalintensität zunimmt, ist zu vermuten, dass dieser Effekt bei realen Messungen geringer ist. Bei realen Messungen ist die Lage des Schwellenwerts zudem nach oben hin durch die Übergänge der Signalkurven in die Plateauphasen beschränkt und muss aufgrund deren Existenz sogar möglichst niedrig gewählt werden, wie in **3.1.3** noch gezeigt wird. Eine größere Effizienz bewirkt eine geringere Streuung der C_T-Werte, weil die log-linearen Abschnitte der Signalkurven steiler sind. Gleichzeitig wird aber die Steigung der Kalibriergeraden (**Gl. 1.7**) flacher, so dass kleinere Abweichungen der C_T-Werte größere Abweichungen der berechneten initialen Templatmengen zur Folge haben. Dies bewirkt letztlich eine Zunahme der Streuung der berechneten initialen Templatmengen (CV) bei zunehmender Effizienz, die im Rahmen praktisch erreichter Effizienzen (1.7-2.0) jedoch vernachlässigbar gering ist (von 10 % auf 11.6 %).

3.1.3 Einfluss der Plateauphase auf die Quantifizierung

Betrachtet man die frühe, exponentielle Phase der PCR, so lässt sich die Produktentwicklung unter Vernachlässigung limitierender Faktoren nach Gleichung **Gl. 1.4** durch eine Exponentialfunktion beschreiben, woraus sich die lineare Abhängigkeit zwischen dem Logarithmus der initialen Templatkopienzahl N_0 und dem Schwellenwertzyklus C_T ableiten lässt (siehe **Gl. 1.7**). Tatsächlich handelt es sich bei der Produktzunahme jedoch um begrenzt exponentielles Wachstum, was sich durch die folgende Gleichung beschreiben lässt:

$$N_c = \frac{N_0 \cdot \varepsilon^c}{1 + N_0 \cdot \varepsilon^{c-m}} \qquad \textbf{Gl. 3.1}$$

Die zugehörige Kurve steigt initial exponentiell an (mit ε als Effizienzparameter), durchläuft einen Wendepunkt und nähert sich für große Zyklenzahlen c einem Plateauwert. Der Parameter m bestimmt hier die Lage des Wendepunktes der sigmoiden Kurve. Der Plateauwert P ergibt sich unter den Annahmen $N_0 \geq 0$ und $\varepsilon \geq 1$ durch

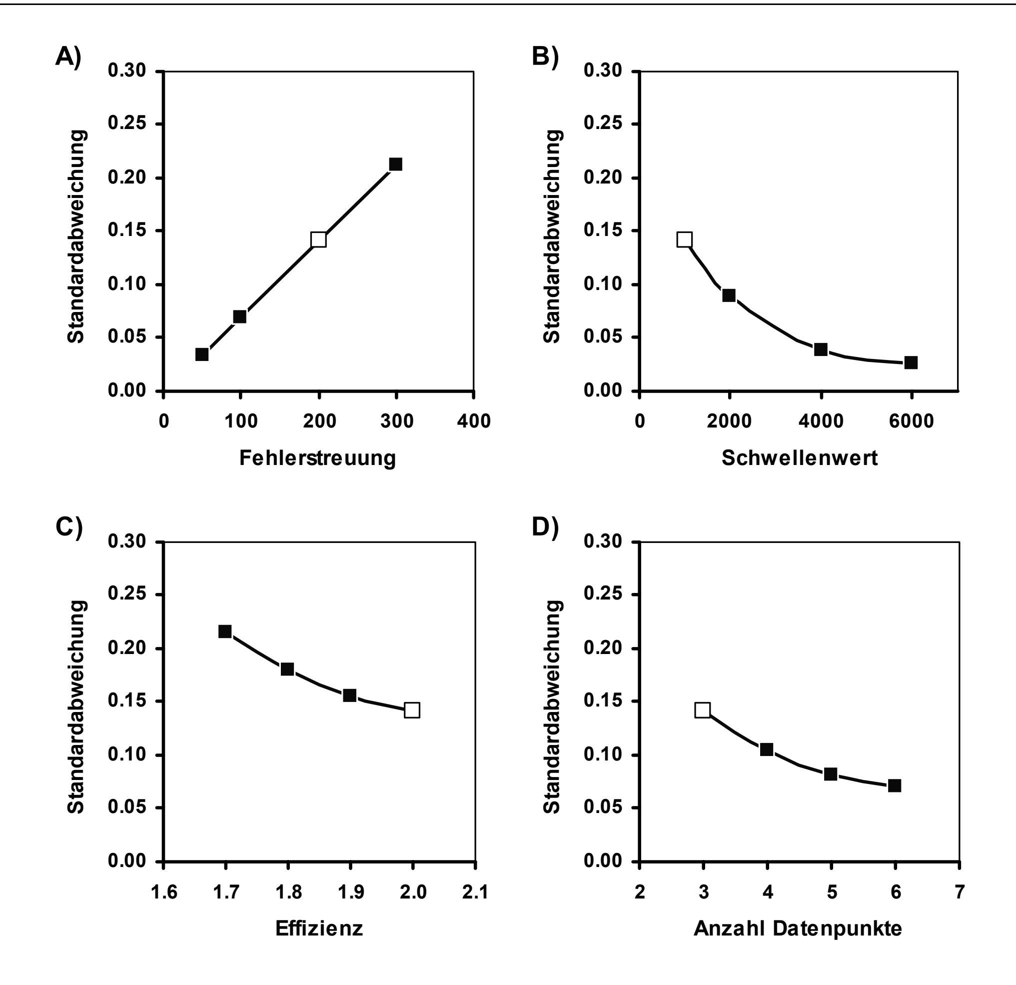


Abb. 3-3 **Streuung der C_T-Werte in Abhängigkeit von verschiedenen Parametern.** Gezeigt sind die SD der **C_T-Werte** [in Zyklen] (n = 1000). Die C_T-Werte wurden anhand simulierter, exponentieller Signalkurven berechnet. Die Werte der Datenpunkte dieser Kurven sind berechnet nach der Formel $y_i = N_0 \cdot \varepsilon^i + \mathrm{err}(f)$, wobei i die Zyklusnummer bezeichnet und die Funktion err einen normalverteilten Zufallswert mit einem MW von Null und einer SD von f zurückgibt.[210] Die Parameter sind, sofern nicht anders angegeben: Initiale Templatmenge $N_0 = 1000$, Effizienz $\varepsilon = 2.0$, Signalstreuung $f = 200$, Schwellenwert $S = 1000$, Anzahl der Datenpunkte für die Regressionsanalyse = 3. Diese Werte gelten immer für den weiß markierten Wert (□) in den Diagrammen. Für die Regressionsanalyse wurden die logarithmierten Signalwerte $\log(y_i)$ verwendet. Anhand der Parameter der Regressionsgeraden (Steigung m und Achsenabschnitt b) sind die C_T-Werte berechnet als $C_T = (\log(S) - b) \cdot m^{-1}$ (vom Schnittpunkt der Graden mit dem Schwellenwert $m \cdot C_T + b = \log(S)$, aufgelöst nach C_T). Die initialen Templatmengen sind berechnet durch $S \cdot \varepsilon^{-C_T}$ (aus $y_i = N_0 \cdot \varepsilon^i$ folgt bei $i = C_T$: $y_i = S$. Damit ist $S = N_0 \cdot \varepsilon^{C_T}$. Aufgelöst nach N_0 ergibt sich $N_0 = S \cdot \varepsilon^{-C_T}$). Die Generierung der Datenpunkte mit anschließender Berechnung der C_T-Werte wurde je Parametersatz 1000mal wiederholt, wobei folgende Parameter variiert wurden: **A)** Fehlerstreuung (f), **B)** Schwellenwert (S), **C)** Effizienz (ε), **D)** Anzahl zur Anpassung verwendeter Datenpunkte (n).

$$P = \lim_{x \to \infty} \frac{N_0 \cdot \varepsilon^c}{1 + N_0 \cdot \varepsilon^{c-m}} = \varepsilon^m \qquad \textbf{Gl. 3.2}$$

Für den Parameter m folgt damit in Abhängigkeit eines gegebenen Plateauwertes die Beziehung $m = \log(P) \cdot (\log(\varepsilon))^{-1}$. Eingesetzt in **Gl. 3.1** ergibt sich nach Vereinfachung

$$N_c = \frac{P \cdot N_0 \cdot \varepsilon^c}{P + N_0 \cdot \varepsilon^c} \qquad \textbf{Gl. 3.3}$$

Mit einer zur Produktmenge N proportionalen Signalintensität S ($S = p \cdot N$), wie es z. B. für die Verwendung eines hohen Überschusses an SYBR-Green I als Farbstoff gegeben ist, folgt für den Schwellenwertzyklus $c = C_T$

$$\frac{1}{p} \cdot S_{C_T} = \frac{P \cdot N_0 \cdot \varepsilon^{C_T}}{P + N_0 \cdot \varepsilon^{C_T}} \qquad \textbf{Gl. 3.4}$$

Aufgelöst nach dem Schwellenwertzyklus ergibt sich daraus

$$C_T = -\frac{1}{\log(\varepsilon)} \cdot \log(N_0) - \frac{\log\left(\frac{P \cdot p - S_{C_T}}{P \cdot S_{C_T}}\right)}{\log \varepsilon} \qquad \textbf{Gl. 3.5}$$

Mit $m = -(\log(\varepsilon))^{-1}$ und $b = m \cdot \log\left(\frac{P \cdot p - S_{C_T}}{P \cdot S_{C_T}}\right)$ ergibt sich wiederum die einfache Geradengleichung (vgl. **Gl. 1.7**), wobei hier der Achsenabschnitt b allerdings vom nicht notwendigerweise für alle Proben gleichen Plateauwert P abhängt. Formuliert man den Logarithmus des Parameters b als Summe

$$b = m \cdot (\log(p \cdot P - S) - \log(S) - \log(P))$$

so erkennt man an diesem Ausdruck, dass sich der Term $\log(p \cdot P - S)$ für $p \cdot P \gg S$ (bzw. für $S \to 0$; ein im Vergleich zum Plateauwert sehr niedriger Schwellenwert, der zur Bestimmung der C_T–Werte dient) näherungsweise durch $\log(p \cdot P)$ bzw. $\log(p) + \log(P)$ beschreiben lässt, womit sich b zu $\log(p) - \log(S)$ vereinfacht. Unter der Voraussetzung, dass der Schwellenwert zur C_T-Wert-Bestimmung deutlich niedriger ist als der Plateauwert, kann der Einfluss des Plateauwertes auf den C_T-Wert bei begrenzt exponentiellem Wachstum vernachlässigt werden und somit die Konzentrationsbestimmung nach der Schwellenwertmethode erfolgen. Ein Beispiel für die quantitative Auswirkung unterschiedlicher Plateauwerte auf die C_T-Werte ist in **Abb. 3-4** dargestellt.

Bei üblichen Parameterwerten und einer Variation der Plateauwerte um einen Faktor 2 liegt die dadurch bedingte Variation der C_T-Werte in einer Größenordung von < 0.1 Zyklen, was eine Fehlbestimmung von maximal $2^{0.1}$-1 ≈ 0.07 bzw. 7 % zur Folge hat.

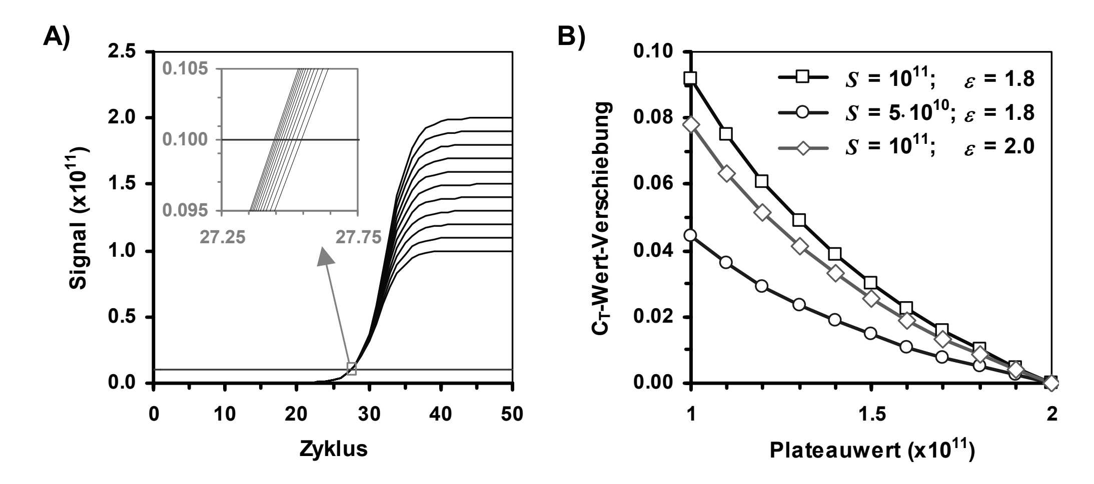


Abb. 3-4 Auswirkung unterschiedlicher Plateauwerte auf die C_T-Werte. A) Berechnete Amplifikationskurven. Funktion siehe Gl. 3.3 mit $N_0 = 10^3$ K; $\varepsilon = 1.8$; $P = 2{\cdot}10^{11}$ K; Schwellenwert $S = 10^{10}$ K (horizontale Linie). Die Ausschnittsvergrößerung zeigt den Intersektionsabschnitt des Schwellenwerts. **B)** Relative Lage des C_T-Werts in Abhängigkeit des Plateauwerts bei unterschiedlichen Schwellenwerten und Effizienzen. Die Verschiebung bezieht sich auf den C_T-Wert der Kurve mit einem Plateauwert von $2{\cdot}10^{11}$ K. Berechnung der C_T-Werte nach Gl. 3.5. Die blaue Kurve (–□–) repräsentiert die Daten aus A).

Für die Verwendung von Hybridisierungssonden kann das Modell um das Massenwirkungsgesetz erweitert werden. In diesem Fall ist das Signal im Zyklus c proportional zu dem im Ansatz befindlichen Komplex Q_c aus Sonde und Produkt. Mit F als der Gesamtmenge der im Ansatz befindlichen Sonde, F' der freien Sonde, N_c' dem freien Produkts im Zyklus c und K als Gleichgewichtskonstante für die Komplexbildung folgt aus dem Massenwirkungsgesetz für $N_c' + F' \rightleftharpoons Q_c$ die Beziehung

$$K = \frac{Q_c}{F' \cdot N_c'} \qquad \textbf{Gl. 3.6}$$

Nach Substitution der Nebenbedingungen ($F = F' + Q_c$ und $N_c = N_c' + Q_c$) ergibt sich

$$K = \frac{Q_c}{(F - Q_c)\cdot(N_c - Q_c)} \qquad \textbf{Gl. 3.7}$$

Für Q_c folgt aus **Gl. 3.7**

$$Q_c = \tfrac{1}{2}\cdot\left(K + F + N_c - \sqrt{(K + F + N_c)^2 - 4\cdot F\cdot N_c}\right) \qquad \textbf{Gl. 3.8}$$

wobei sich N_c bei begrenzter exponentieller Produktzunahme aus **Gl. 3.1** ergibt.

Setzt man analog zum oben beschriebenen Verfahren (vergleiche **Gl. 3.3**) als Signalwert den Schwellenwert ein ($S_{C_T} = p \cdot Q_{C_T}$ mit p als Proportionalitätsfaktor) und löst die Gleichung nach dem Schwellenwertzyklus C_T auf, erhält man

$$C_T = -\frac{1}{\log \varepsilon} \cdot \log(N_0) - \frac{\log\left(\frac{{S_{C_T}}^2 + F \cdot P \cdot p^2 - S_{C_T} \cdot p \cdot (K + F + P)}{P \cdot S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)}\right)}{\log \varepsilon} \qquad \text{Gl. 3.9}$$

Als Parameter der Kalibriergeraden ergeben sich hieraus

$$m = -\left(\log(\varepsilon)\right)^{-1} \text{ und } b = m \cdot \log\left(\frac{{S_{C_T}}^2 + F \cdot P \cdot p^2 - S_{C_T} \cdot p \cdot (K + F + P)}{P \cdot S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)}\right)$$

Auch hier ist der Parameter b wieder vom Plateauwert abhängig. Unter der Voraussetzung, dass der Schwellenwert sehr viel kleiner ist als der Plateauwert, kann diese Abhängigkeit jedoch vernachlässigt werden. Zum Nachweis setze man $P = r \cdot S_{C_T}$, um das Plateausignal als Vielfaches des Schwellenwertsignals zu beschreiben und löse den Grenzwert des Arguments des Logarithmus für r gegen unendlich:

$$\lim_{r\to\infty}\left(\frac{S_{C_T} \cdot \left(S_{C_T} + r \cdot p^2 \cdot F - p \cdot \left(K + F + r \cdot S_{C_T}\right)\right)}{r \cdot {S_{C_T}}^2 \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)}\right) = \lim_{r\to\infty}\left(\frac{r \cdot \left(p^2 \cdot F - p \cdot S_{C_T}\right) + S_{C_T} - p \cdot (K + F)}{r \cdot S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)}\right) =$$

$$\lim_{r\to\infty}\left(\frac{r \cdot \left(p^2 \cdot F - p \cdot S_{C_T}\right)}{r \cdot S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)} + \frac{S_{C_T} - p \cdot (K + F)}{r \cdot S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)}\right) =$$

$$\frac{\left(p^2 \cdot F - p \cdot S_{C_T}\right)}{S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)} + \lim_{r\to\infty}\left(\frac{1}{r}\right) \cdot \frac{S_{C_T} - p \cdot (K + F)}{S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)} = \frac{p^2 \cdot F - p \cdot S_{C_T}}{S_{C_T} \cdot \left(p \cdot K + p \cdot F - S_{C_T}\right)}$$

Der Plateauwert ist nach der Grenzwertbildung nicht mehr Teil des Arguments, womit der C_T-Wert formell nur noch von den Konstanten p, F und K sowie vom Schwellenwert S_{C_T} abhängt.

3.1.4 Einfluss hochschmelzender Regionen genomischer Zielsequenzen

Während Plateauwert-bedingte Abweichungen der C_T-Werte vernachlässigbar sind (siehe **3.1.2**), haben Schwankungen der Amplifikationseffizienzen einen sehr großen Einfluss auf den C_T-Wert. Derartige Schwankungen auf die Effizienz können statistischer Natur sein, die sich insbesondere bei sehr geringen initialen Templatmengen bemerkbar machen, oder durch Unregelmäßigkeiten in den Reaktionsbedingungen bedingt sein. Durch eine mög-

lichst gute Gleichbehandlung aller Reaktionsansätze und die Verwendung von *master*-Reaktionsansätzen wo immer möglich (vgl. **2.6.1**), können derartige Unregelmäßigkeiten auf ein Minimum begrenzt werden. Es gibt jedoch einen wichtigen Parameter, auf den der Experimentator keinen Einfluss hat: die Temperaturhomogenität im PCR-Gerät.

Über die Temperatur werden bei der PCR die Denaturierung, Primer/Templat-Hybridisierung und Polymerisierung gesteuert. Besonders bei *rapid-cycle*-PCRs kann sich eine moderate Abweichung der Temperatur bei jedem dieser Schritte merklich auf die Effizienz und damit auf die C_T-Werte auswirken. Während die Temperaturen für Hybridisierung und Polymerisierung gut an die jeweiligen Erfordernisse angepasst werden können, ist der Wahl für die Denaturierungstemperatur eine obere Grenze gesetzt. Zwar denaturiert hochmolekulare DNA im Mittel zwischen 90-95 °C, dennoch gibt es stabile, G:C-reiche Inseln im

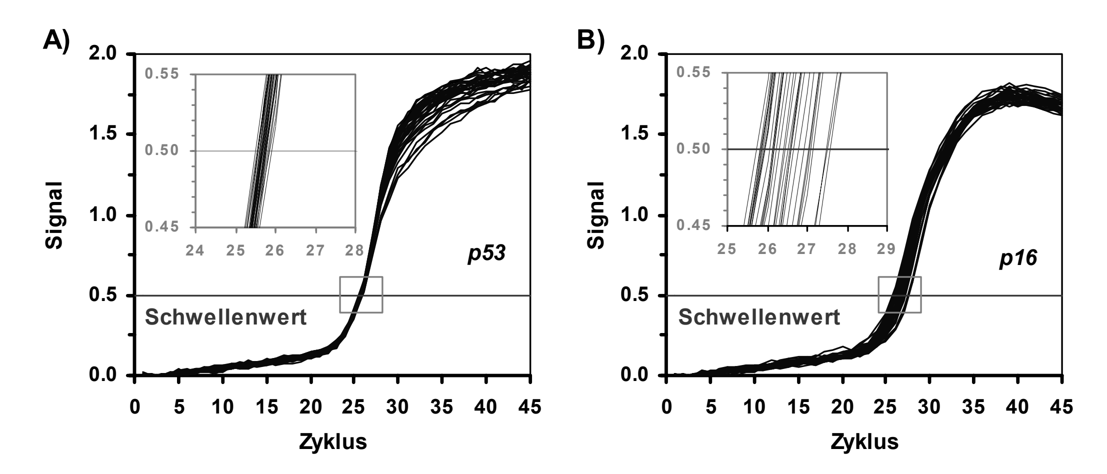


Abb. 3-5 Amplifikationskurven von Mehrfachansätzen und resultierende Verteilungen der C_T-Werte. Je ein mit 32 identischen PCR-Ansätzen mit 3 ng humaner genomischer DNA als Templat bestücktes Probenkarussell wurde dem Standardtemperaturprotokoll im LightCycler unterzogen. Die Detektion erfolgte mit Hybridisierungssonden. Die aufgetragenen Signalwerte sind das Verhältnis von Kanal 1 zu Kanal 2. Die Ausschnittsvergrößerungen zeigen die Signalkurven im Bereich um die C_T-Werte. **A)** Zielsequenz *p53*. Mittlerer C_T-Wert: 25.7 ± 0.07. **B)** Zielsequenz *p16*. Mittlerer C_T-Wert: 26.4 ± 0.55.

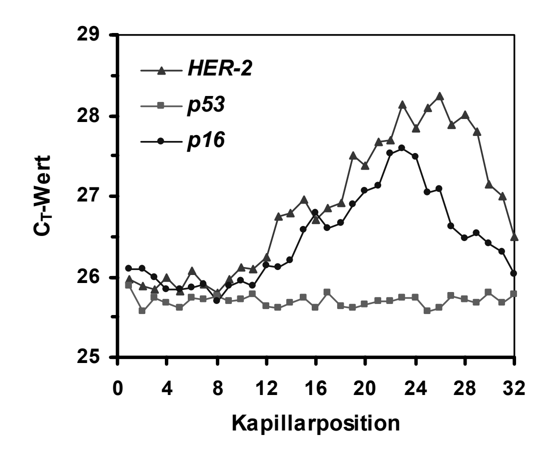


Abb. 3-6 C_T-Werte in Abhängigkeit von der Position der Kapillare im LightCycler Probenkarussell. Durchführung und Auswertung sind der Legende von Abb. 3-5 zu entnehmen. Für die HER-2-Zielsequenz wurden die Primer HER-H1/R1 und die Sonden HER-D1/A1 verwendet.

Genom, die auch bei 95°C nicht vollständig aufschmelzen. Nach dem Abkühlen auf die Hybridisierungstemperatur können die nur partiell aufgeschmolzenen Templatstränge sehr schnell rehybridisieren und so eine Anlagerung der Primer behindern. Eine geringfügig niedrigere Temperatur kann sich hier sehr stark auswirken. Denaturierungstemperaturen oberhalb von 95 °C sind hingegen kritisch, da bei diesen Temperaturen zum einen die Halbwertszeit der meist verwendeten thermostabilen *Taq* DNA-Polymerasen überproportional stark abnimmt.

Temperaturinhomogenitäten im PCR-Gerät haben also besonders großen Einfluss auf die Effizienzen während der Denaturierung in den ersten Zyklen der PCR, wo die Menge hochmolekularen Templats noch signifikant ist im Vergleich zur Menge an PCR-Produkten, wenn sich die Primerbindungsstellen in der Nähe von hochschmelzenden Regionen befinden. Dieser Einfluss wird im Folgenden gezeigt und quantitativ untersucht. In Mehrfachansätzen erwartet man identische C_T-Werte für alle Reaktionsansätze, was in vielen Fällen auch beobachtet wird (**Abb. 3-5 A**). Manche Primerkombinationen (p16-H/R, HER-H1/R1) zeigen jedoch unter sonst gleichen Reaktionsbedingungen im selben PCR-Gerät sehr große Schwankungen der C_T-Werte (**Abb. 3-5 B**).

Die Verteilung der C_T-Werte ist hierbei jedoch nicht zufällig bzw. normalverteilt. Sie zeigt bei anfälligen Zielsequenzen eine deutliche Abhängigkeit von der Position des Reaktionsansatzes im PCR-Gerät (**Abb. 3-6**).

Die gegen die Kapillarposition im LightCycler aufgetragenen C_T-Werte zeigen einen sinusoidalen Verlauf als impliziter Hinweis auf die rotationssymmetrische Anordnung der Kapillaren. Diese typische Abhängigkeit der C_T-Werte von der Kapillarposition wurde für diese Primerkombinationen an drei von insgesamt sieben unterschiedlichen getesteten LightCycler-Geräten an verschiedenen Orten beobachtet. Die Phase und Amplitude dieser Verteilung wurde hinsichtlich des Einflusses verschiedener Parameter untersucht (**Abb. 3-7**).

Das Phänomen ist dabei unabhängig von der Quelle der DNA, die von unterschiedlichen Personen und Isolierungen stammte, dem Detektionsformat (SYBR-Green I oder Hybridisierungssonden), der Hybridisierungszeit (**Abb. 3-7 A**) und -temperatur (**Abb. 3-7 B**), der Denaturierungszeit (0 s, 5 s und 10 s bei 95 °C) sowie der unmittelbaren Primersequenz, wie die vergleichende Untersuchung verschiedener Primerpaare zeigte, deren Bindungsstellen sich alle in der unmittelbaren Nähe des HER-H1/R2-Primerpaares befinden. Eine Erhöhung der Denaturierungstemperatur verringert die Amplitude der positionsabhängigen Schwankungen der C_T-Werte (**Abb. 3-7 C**). Die Schwankungen treten nicht auf, wenn statt humaner genomischer DNA aufgereinigtes PCR-Produkt oder mit einem Restriktionsenzym in der Nähe der Amplikonsequenz gespaltene genomische DNA als Templat verwendet wird (**Abb. 3-7 D**). Die Streuung der C_T-Werte bei der HER-2-Zielsequenz tritt für ein weiteres, Primerpaar, HER-2-H2/R2 mit den Sonden HER-2-D2/A2, nicht auf. Die Bindungsstellen dieses Primerpaares liegen etwa 1 kBp weiter im Gen.

Anhand der Amplitude der sinusoidalen Schwankungen der C_T-Werte (ca. 1.5 Zyklen bei einer Denaturierungstemperatur von 95 °C, siehe **Abb. 3-6 C**) und der Abhängigkeit der mittleren C_T-Werte von der Denaturierungstemperatur (ca. -0.5 Zyklen/K, siehe **Abb. 3-8**)

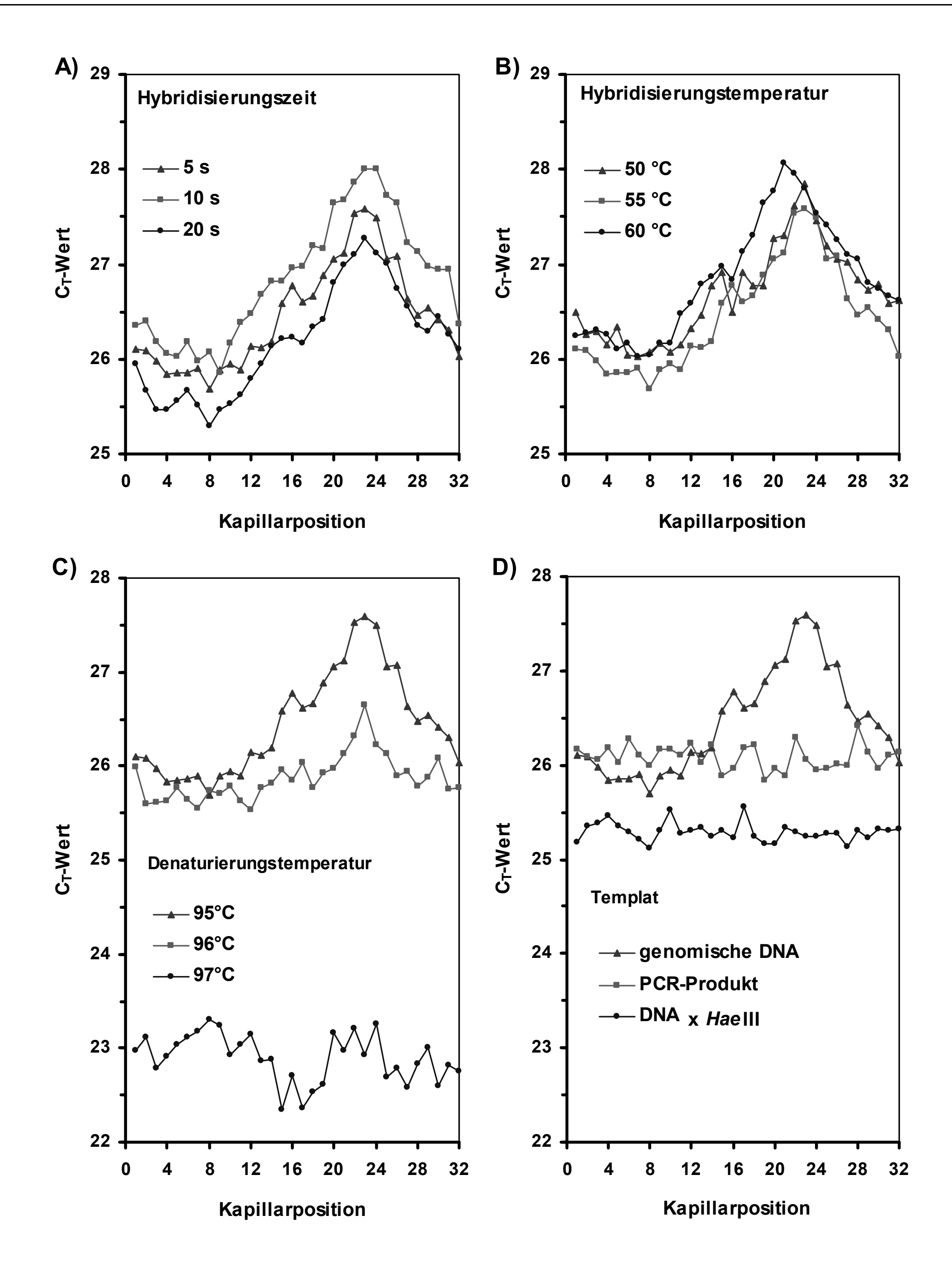


Abb. 3-7 **C_T-Werte in Abhängigkeit von der Kapillarposition im LightCycler.** Die C_T-Werte für die *p16*-Zielsequenz wurden ermittelt wie in Abb. 3-6 beschrieben. Einfluss von **A)** Hybridisierungsdauer, **B)** Hybridisierungstemperatur, **C)** Denaturierungstemperatur und **D)** der Zielsequenz benachbarter genomischer Bereiche. Das 298 Bp lange *Hae*III-Restriktionsfragment genomischer DNA enthält die *p16*-Zielsequenz.

lässt sich abschätzen, dass die maximale Temperaturdifferenz in der Probenkammer ca. ±1.5 K beträgt, was gut übereinstimmt mit den vom Hersteller angegebenen technischen Spezifikationen des LightCyclers, der eine Temperaturpräzision von ±1.0 K bei 95 °C nennt (LightCycler Benutzerhandbuch, Version 1.2).

Bei einer Amplifikationseffizienz von 1.9 beträgt der relative Fehler in der Quantifizierung für die hier gezeigten Beispiele $1.9^{1.5} - 1 \approx 1.62$ bzw. 162 %. Dieser Fehler lässt sich aus der Kalibriergeraden oder aus der Streuung der C_T-Werte von Mehrfachansätzen mit $n < 15$ nur sehr bedingt ableiten. So ist in den gezeigten Beispielen etwa die Streuung der C_T-Werte für die Proben in den ersten 10 Kapillarpositionen bei allen Zielsequenzen und unter allen Versuchsbedingungen vergleichbar (siehe **Abb. 3-6 C**, **Abb. 3-7** und **Abb. 3-8**).

In der Summe zeigen die Ergebnisse, dass Temperaturschwankungen während der Denaturierung hochmolekularer Template in den ersten Zyklen der PCR einen entscheidenden Einfluss auf die Lage der C_T-Werte und damit auf die Ergebnisse von Quantifizierungen haben können. Die Zielsequenzen der Primer (p16-H/R und HER-H1/R1), welche für derartige Schwankungen bzw. Inhomogenitäten anfällig waren, zeigen in ihrer Nähe auf den genomischen Templaten hochschmelzende Regionen.

Mit der Software Melt94 wurden für beide anfälligen Zielsequenzen in etwa 100 Bp Entfernung von den Primerbindungsstellen Regionen identifiziert, deren berechnete Schmelztemperatur 10-15 K höher war, als die der übrigen Sequenz. Für die getesteten nicht-anfälligen Zielsequenzen (p53, HBB, IGF-1) sowie die genomische Region um die Bindungsstellen der Primer HER-2-H2/R2 konnten in einer Entfernung bis 500 Bp von den Primerbindungsstellen keine hochschmelzenden Regionen identifiziert werden.

Der mögliche Fehler bei der Quantifizierung anfälliger Zielsequenzen kann minimiert werden, indem die genomische Templat-DNA vor der PCR für einige Minuten im Wasserbad gekocht wird, um sie initial möglichst vollständig zu denaturieren. Eine Verbesserung der Präzision ergibt sich auch, wenn sich die Kapillaren nicht während der gesamten Denaturierungszeit am selben Ort im LightCycler befinden, sondern durch die Probenkammer bewegt werden. Dies kann bei LightCycler-Geräten aktueller Baureihen durch eine kontinuierliche Rotation des Probenkarussells erreicht werden.

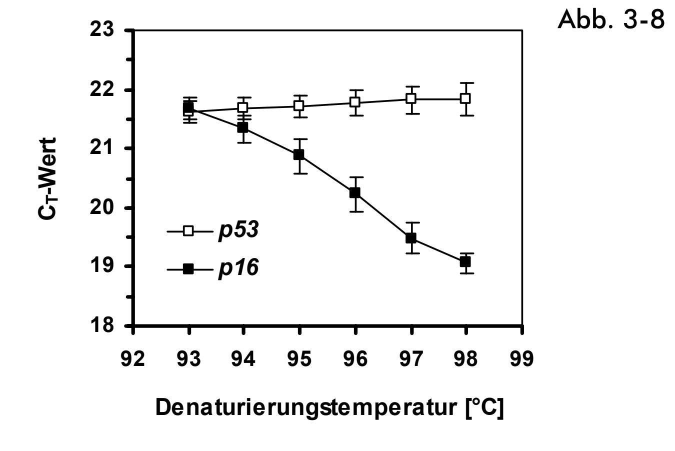


Abb. 3-8 Mittlere C_T-Werte in Abhängigkeit von der Denaturierungstemperatur in den initialen Zyklen. Ausgewertet wurden die C_T-Werte der ersten 10 Kapillarpositionen wie in Abb. 3-6 beschrieben (MW±SD) von 6 Experimenten mit der jeweils angegebenen Denaturierungstemperatur in den ersten 10 Zyklen des Experiments. In den folgenden Zyklen betrug die Denaturierungstemperatur immer 95 °C. Die Templatmenge betrug hier 30 ng je Reaktionsansatz.

3.1.5 Einfluss des Sondenabbaus auf Quantifizierungen

Zur sequenzspezifischen Generierung von Fluoreszenzsignalen während der PCR stehen verschiedene Formate zur Verfügung (vgl. **0**). Bei nahezu allen Formaten führt die Hybridisierung der Sonden mit dem Templat unmittelbar zum Anstieg der Reporter- bzw. Akzeptorfluoreszenz. Die Ausnahme bildet das TaqMan-Format, wo die Sonden von der *Taq* DNA-Polymerase hydrolysiert werden müssen, um den Reporter vom Quencher zu trennen.

Dazu verdrängt die Polymerase zunächst das im Weg befindliche Oligonukleotid partiell und formt so eine Y-förmige Struktur. Das nicht-gepaarte 5'-Ende dieser Verzweigungsstruktur wird dann durch die strukturspezifische Endonukleaseaktivität der *Taq* DNA-Polymerase hydrolysiert.[211,212] Bei längeren Fragmenten wiederholt sich der Vorgang so oft, bis das verbliebene Fragment von der Polymerase einfach vom Templatstrang abgeschert wird. Im Gegensatz dazu sollten die Sonden der anderen Formate gerade nicht hydrolysiert

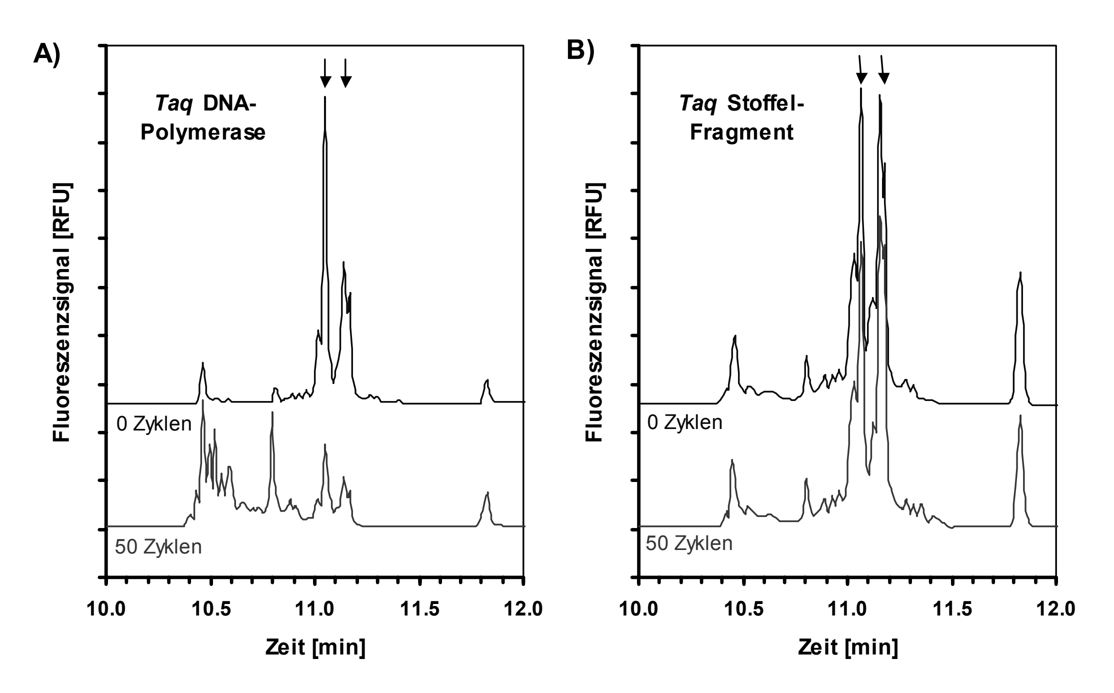


Abb. 3-9 Elektropherogramme der Donorsonden. Mit einer kapillarelektrophoretischen Analyse wurden die von den Donorsonden stammenden, mit Fluoreszein markierten Fragmente aus PCR-Ansätzen untersucht. Aliquots der Reaktionsansätze unmittelbar vor der PCR (obere Kurve) und nach 50 Amplifikationszyklen (untere Kurve) wurden verglichen. Gezeigt sind die Daten für die Zielsequenz *p16*. Die beiden mit Pfeilen markierten Peaks entsprechen den Fragmenten der intakten, nichthydrolysierten Sonden mit den Isomeren 5- und 6-Carboxyfluoreszein. Die Länge der Donorsonde ist 30 nt. Die Peaks linksseitig der Pfeilmarkierung kennzeichnen verkürzte fluoreszeinmarkierte Fragmente. Die relativen Mengen der Fragmente wurden anhand der Peakflächenintegrale bestimmt. Die mit RED-640 markierte Akzeptorsonde wurde nicht detektiert. **A)** Fragmente der Donorsonde vor und nach Amplifikation mit *Taq* DNA-Polymerase. **B)** Fragmente der Donorsonde vor und nach Amplifikation mit dem Stoffel-Fragment der *Taq* DNA-Polymerase.

werden, um die effektive Sondenkonzentration während der PCR nicht zu verringern. Hydrolysierte, unmarkierte Fragmente könnten zudem mitunter mit den intakten Sonden in der Hybridisierung konkurrieren, aber nicht zum Signal beitragen. Interessanterweise wird in allen Formaten die gleiche *Taq* DNA-Polymerase (mit Endonukleasedomäne) benutzt wie im TaqMan-Format. Tatsächlich werden die im LightCycler häufig verwendeten Hybridisierungssonden ebenfalls während der PCR durch die Polymerase hydrolysiert, wie durch die kapillarelektrophoretische Analyse der FAM-markierten Donorsonden gezeigt werden kann (**Abb. 3-9 A**).

Im Anschluss an die PCR befindet sich ein deutlich größerer Anteil sehr kurzer, FAM-markierter Fragmente in den Reaktionsansätzen als vor der PCR. Bei den kurzen Fragmenten kann es sich nur um die Hydrolyseprodukte der Sonden handeln. Die Sondenhydrolyse wurde mit unterschiedlichen Primer/Sonden-Systemen untersucht (Zielsequenzen *p16*, *p53*, *IGF-1*, *HBB* und *bcr/abl*). In allen untersuchten Fällen findet während der PCR eine Hydrolyse der Sonden statt, wobei das Ausmaß jedoch unterschiedlich ist (zwischen 20 und 70 %, berechnet anhand der Peakflächenintegrale in den Elektropherogrammen) und vermutlich vom Sequenzkontext abhängt. Wird statt der üblichen *Taq* DNA-Polymerase das Stoffel-Fragment, eine Variante der *Taq* DNA-Polymerase (ohne Endonukleasedomäne), verwendet, so tritt kein erkennbarer Sondenabbau auf (**Abb. 3-9 B**). Dies zeigt, dass der Sondenabbau nur durch die endonukleolytische Aktivität der *Taq* DNA-Polymerase bedingt ist. Dabei entstehen nicht Fragmente beliebiger Längen, wie es bei einem sequenziellen

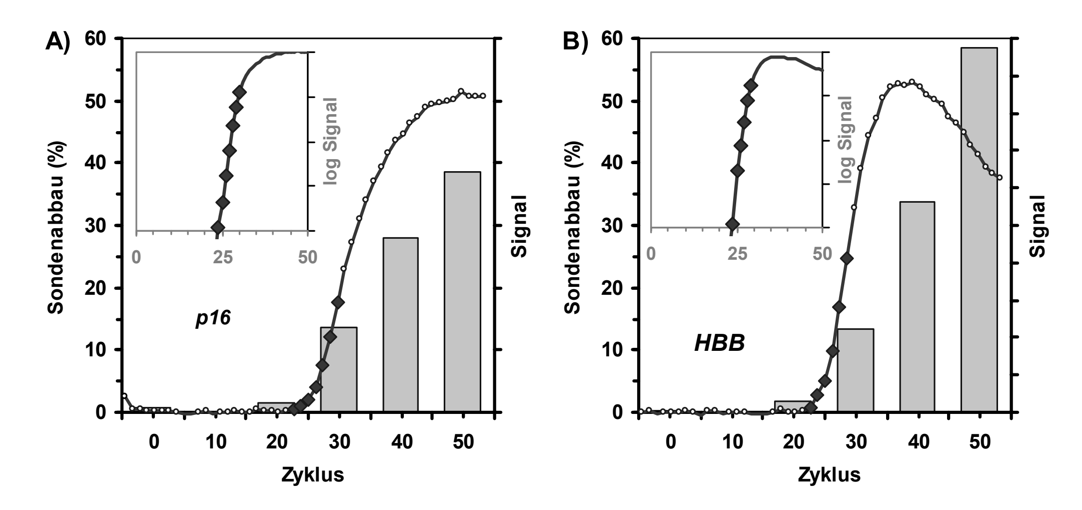


Abb. 3-10 Kinetik der Sondenhydrolyse. Aliquots von sechs Replikaten eines Reaktionsansatzes mit *Taq* DNA-Polymerase wurden nach unterschiedlicher Anzahl Amplifikationszyklen kapillarelektrophoretisch untersucht. Gezeigt sind die Anteile verkürzter Fragmente (blaue Balken), die sich berechnen nach $A_H/A_T - A_{Ho}/A_{To}$, wobei A_H: Summe der Peakflächen der verkürzten Fragmente, A_T: Summe aller Peakflächen, A_{Ho}: A_H vor PCR, A_{To}: A_T vor PCR. Den Daten sind die Amplifikationssignale während der PCR überlagert (Kanal 2 / Kanal 1: rote Kurve). Die Punkte in den exponentiellen Phasen sind durch rote Rhomben markiert. Die eingefügten Diagramme zeigen das Amplifikationssignal in halblogarithmischer Auftragung. **A)** Zielsequenz *p16*. **B)** Zielsequenz *HBB*.

Abbau der Sonden zu erwarten stünde. Vielmehr entstehen bevorzugt Fragmente, die etwa 15 Nukleotide kürzer sind als die vollständigen Sonden, während nur um wenige Nukleotide verkürzte Fragmente praktisch nicht detektierbar sind.

Dieser Befund ist bei strukturspezifischen Endonukleasen zu erwarten und bestätigt die bereits beschriebenen Eigenschaften der *Taq* DNA-Polymerase.[211,212] Der Anteil hydrolysierter Sonden nimmt während der PCR ungefähr parallel mit der Produktmenge zu (**Abb. 3-10**). Detektierbare Mengen verkürzter Fragmente treten erst in der exponentiellen Phase auf und erreichen gegen Ende der exponentiellen Phase einen Anteil von bis zu 15 %. Bei Verwendung des Stoffel-Fragments erreichen die Amplifikationskurven wesentlich höhere Signalwerte als bei Verwendung der *Taq* DNA-Polymerase (**Abb. 3-11 A**). Dadurch werden das Signal/Rausch-Verhältnis sowie die Anzahl der Datenpunkte im Bereich des exponentiellen Signalanstiegs und mithin die Präzision der C_T-Wertbestimmung erhöht. Versuche mit Mehrfachansätzen (n=8) verschiedener Zielsequenzen (*p16*, *p53*, *IGF-1*, *HBB*) zeigen, dass die SD der C_T-Werte der mit dem Stoffel-Fragment amplifizierten Proben immer 20-30 % geringer ist als die der mit *Taq* DNA-Polymerase amplifizierten Proben.

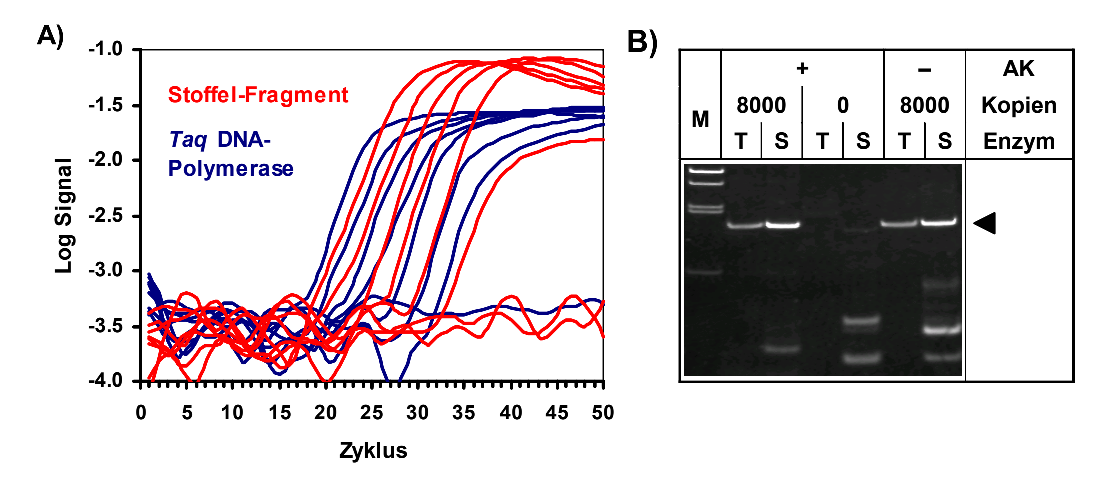


Abb. 3-11 **Vergleich von *Taq* DNA-Polymerase und Stoffel-Fragment. A)** Signalkurven der Hybridisierungssonden nach Amplifikation der *IGF-1*-Zielsequenz mit den Primern IGF-H1/R1 aus humaner genomischer DNA (8000, 2000, 500, 125, 32, 8, 2, 0 Kopien) mit *Taq* DNA-Polymerase (blau) oder Stoffel-Fragment (rot). Die Detektion erfolgte mit den *IGF-1*-spezifischen Hybridisierungssonden. Die Kurven zeigen die exponentiellen Anstiege in der Reihenfolge abnehmender Templatmengen von links nach rechts. Die Kurven der Negativkontrollen (0 Kopien) zeigen keinen Anstieg. Beim Stoffel-Fragment zeigt zusätzlich auch die Kurve der Probe mit 2 Kopien keinen Signalanstieg. **B)** PAGE-Analyse der PCR-Produkte aus A). M: Fragmentlängenstandard pTR54 x *Hin*fI (oberste Bande: 220/221 Bp), AK: Antikörper, T: *Taq* DNA-Polymerase, S: Stoffel-Fragment. Der Pfeil kennzeichnet die Position des spezifischen IGF-1-H/R-Produkts (123 Bp). Die Amplifikate mit hoher elekrophoretischer Mobilität stellen unspezifische Nebenprodukte dar.

Der höheren Präzision bei Verwendung des Stoffel-Fragments stehen allerdings eine geringere Sensitivität und Genauigkeit gegenüber. Weniger als 30 initiale Templatkopien lassen sich mit dem Stoffel-Fragment nur selten nachweisen, während die *Taq* DNA-Polymerase selbst solitäre Templatkopien im Reaktionsansatz gut amplifiziert (statistisch zeigt ca. jede zweite PCR mit nur einer einzigen Templatkopie eine Amplifikation [213]). So erreicht z. B. die Probe mit 8 initialen Templatkopien in **Abb. 3-11 A** bei Amplifikation mit dem Stoffel-Fragment nur noch einen Bruchteil der Signalintensität der Proben mit mehr Templatkopien und die Probe mit 2 initialen Kopien zeigt keine erkennbare Amplifikation mehr. Weiterhin entstehen bei Verwendung des Stoffel-Fragments vergleichsweise große Mengen unspezifischer Nebenprodukte in der PCR (**Abb. 3-11 B**).

Die Bildung dieser Nebenprodukte kann durch die Anwendung einer *Hot-start*-Technik zwar deutlich verringert, aber nicht ganz ausgeschlossen werden. Durch die Bildung von Nebenprodukten werden Primer aufgebraucht und DNA-Polymerase gebunden, wodurch ihre effektive Konzentration zur Amplifikation des spezifischen Produkts und damit auch die Amplifikationseffizienz abnimmt. Da die Entstehung unspezifischer Produkte während der PCR eine statistische Komponente hat, kann sich die Amplifikationseffizienz in den individuellen Ansätzen unterscheiden, was bei der Quantifizierung mit externer Standardisierung zu fehlerhaften Ergebnissen führen kann.

Durch die Verwendung des Stoffelfragments kann die Hydrolyse von Hybridisierungssonden während der PCR verhindert werden, was zu besseren Signalausbeuten führt. Die dadurch längeren exponentiellen Phasen der Amplifikationskurven mit besserem Signal/Rausch-Verhältnis lassen eine höhere Genauigkeit für die Quantifizierung erwarten. Allerdings neigt das Stoffelfragment stark zur Bildung unspezifischer Produkte. Zudem ist die mit dem Stoffelfragment erreichbare Amplifikationseffizienz (und damit die Sensitivität) deutlich geringer als die mit der *Taq* DNA-Polymerase. Während die Bildung unspezifischer Produkte durch den Einsatz einer *Hot-start*-Technik weitgehend unterdrückt werden konnte, entzog sich die Effizienz einer weiteren Optimierung. Zusammenfassend lässt sich sagen, dass die Sondenhydrolyse durch die *Taq* DNA-Polymerase zwar die Qualität der Messergebnisse beeinträchtigt, durch die Verwendung des Stoffelfragments wegen seiner geringen Spezifität und Effizienz allerdings keine Verbesserung erzielt werden kann. Es ist zu vermuten, dass die Verwendung nicht-hydrolysierbarer Sondenanaloga wie z. B. PNAs die Qualität von RT-PCR-Daten zu verbessern vermag.

3.2 Algorithmen und Software

Denn es ist zuletzt doch nur der Geist, der jede Technik lebendig macht.

Johann Wolfgang von Goethe

Neben der Gerätetechnik beeinflussen natürlich auch experimentelle Parameter die erreichbare Genauigkeit von RT-PCR-Experimenten, wie im vorigen Abschnitt gezeigt wurde.

Zur Berechnung der Ergebnisse von RT-PCR-Messungen, sei es zur Quantifizierung oder zur Schmelzkurvenanalyse, müssen die gemessenen Rohdaten prozessiert werden. Die Qualität der Ergebnisse ist also zudem abhängig von der Art der Datenprozessierung, d. h., ausgehend von den Messdaten, von den verwendeten Algorithmen zur Berechnung der Ergebnisse. Die bekannten Algorithmen zur Datenauswertung, implementiert in der mit den RT-PCR-Geräten ausgelieferten Software, nutzen die Informationen der Messdaten nur unzureichend, weshalb die damit ermittelten Ergebnisse bei Messdaten mit geringem Signal-Rausch-Verhältnis oder ungewöhnlichem Verlauf nur eine geringe Präzision aufweisen. Zudem treten unter bestimmten Bedingungen systematische Fehler auf. In den folgenden Abschnitten werden neue Algorithmen zur Auswertung von RT-PCR-Daten vorgestellt, welche nicht nur genauere und präzisere Ergebnisse liefern als die in der Literatur beschriebenen Verfahren, sondern auch eine automatisierte Auswertung einschließlich einer qualitativen Wertung der Ergebnisse erlauben, wie im letzten Teil dieses Abschnitts gezeigt wird, in welchem eine neue Software, welche diese Algorithmen implementiert, vorgestellt wird.

3.2.1 Algorithmus zur quantitativen Auswertung

3.2.1.1 Prinzip

Das Prinzip der RT-PCR mit externer Standardisierung ist in **0** beschrieben. Zunächst werden aus den Amplifikationskurven der PCR-Ansätzen mit Standard- und Proben-DNA deren C_T-Werte bestimmt, dann wird aus den C_T-Werten der Standards eine Kalibrierfunktion erstellt (**Gl. 1.7**), über welche schließlich anhand der C_T-Werte der Proben deren initiale Kopienzahlen berechnet werden. Dieses Verfahren erfordert die Festlegung eines geeigneten Schwellenwerts zur Ermittlung der C_T-Werte, der alle auszuwertenden Amplifikationskurven in ihrer exponentiellen Phase schneidet. Dazu müssen die Signalkurven zunächst um die Hintergrundfluoreszenz korrigiert werden. Im einfachsten Fall handelt es sich dabei um eine Verschiebung, bedingt durch eine unterschiedliche, zeitlich konstante Grundfluoreszenz der Proben, die z. B. durch Abweichungen in der Fluorophormenge oder der Positionierung der Kapillare über dem Linsensystem hervorgerufen sein kann. Häufig lässt sich jedoch schon in den ersten Zyklen der PCR auch ein Signaltrend beobachten, der nicht durch eine Produktakkumulation begründet ist. Diese Ursachen für diese positiven oder negativen Trends sind nicht bekannt, die Stärke der Effekte schwankt mit den Pufferbedingungen und Sondensequenzen. Exzimerbildung und Abspaltung von Fluorophoren können mögliche Ursachen dafür sein. Während sich die Verschiebung sehr einfach korrigieren lässt, erfordert die Korrektur der Signaltrends eine Extrapolation von den ersten Zyklen der PCR ausgehend bis über die exponentielle Phase hinaus, wo der Trend von der amplifikationsbedingten Signalzunahme überlagert wird. Die aktuell von den Herstellern der PCR-Geräte erhältlichen Softwareprogramme zur quantitativen Auswertung von RT-PCR-Daten ermöglichen nur die Korrektur einer unterschiedlichen Hintergrundfluoreszenz, nicht aber eine Trendkorrektur.

Nachdem die Signalkurven um Verschiebungen und amplifikationsunabhängige Signaltrends bereinigt sind, muss beurteilt werden, welche Teile der Signalkurven einen exponentiellen Anstieg aufweisen, um den erlaubten Bereich für den Schwellenwert zu bestimmen. Hierzu werden bei den kommerziell erhältlichen Softwareprogrammen die Signalkurven in halblogarithmischer Auftragung grafisch angezeigt, so dass die exponentiellen Phasen als Geradensegmente erscheinen. Der Anwender muss nun visuell beurteilen, wo er den Schwellenwert definieren möchte. Die Programme der Gerätehersteller machen hierbei keine Aussage über die Gültigkeit der Wahl des Schwellenwertes bzw. dass die zur Berechnung der CT-Werte verwendeten Punkte der Amplifikationskurven auch tatsächlich alle Teil der exponentiellen Phasen sind.

Anschließend werden anhand des Schwellenwertes die C_T-Werte berechnet. Da es sich bei den Daten der Signalkurven um diskrete Werte handelt, müssen die Daten zur Berechnung der C_T-Werte interpoliert werden. Herkömmlicherweise wird dazu der Schwellenwert nicht innerhalb, sondern direkt unterhalb der exponentiellen Phase als „Rauschgrenze“ oder „*noise band*“ gesetzt und eine willkürliche Anzahl – meist 2 bis 3 – von darüber liegenden Punkten im Geradensegment der logarithmierten Kurve für eine Regressionsanalyse verwendet. Der Schnittpunkt der Regressionsgeraden mit dem Schwellenwert ist dann der C_T-Wert.[22] Problematisch bei diesem als *Fit-points*-Methode bezeichneten Ansatz ist, dass die gesetzte Rauschgrenze für alle auszuwertenden Kurven gleichermaßen gilt und oftmals weniger als 4 Datenpunkte je Kurve für die Regressionsanalyse verwendet werden können, wodurch kleine Abweichungen in den Signalwerten einen signifikanten Einfluss auf die Lage der Regressionsgeraden und damit auf den C_T-Wert haben können.

Der in dieser Arbeit entwickelte Algorithmus zur quantitativen Auswertung von RT-PCR-Daten umfasst die Korrektur amplifikationsunabhängiger Signaltrends, die automatische Erkennung der exponentiellen Phasen, die Interpolation der exponentiellen Phasen durch eine Amplifikationsfunktion und die automatische Wahl des optimalen Schwellenwertes.

3.2.1.2 Korrektur amplifikationsunabhängiger Signaltrends

Die Messdaten werden zunächst durch glättende, kubische nichtparametrische *Splines* beschrieben.[214,215] Diese Form der Darstellung der zeitlichen Folge der Messwerte hat den Vorteil, dass die Beziehung zwischen Zyklenzahl und Fluoreszenzsignal durch eine kontinuierliche Funktion in guter Annäherung beschrieben wird. Die Glättung der interpolierenden Funktion ermöglicht die bessere Identifizierung bestimmter Eigenschaften des Signalverlaufs, wobei sie allerdings bei schnellen Signaländerungen, ähnlich wie eine gleitende Mittelwertbildung, eine systematische Abweichung von den Messwerten aufweist. Eine kurze Beschreibung der Herleitung sowie der Eigenschaften kubischer glättender *Splines* ist in **7.2** gegeben.

Anhand der interpolierten kubischen *Splines* wird zunächst der ungefähre Beginn der erkennbaren exponentiellen Signalzunahme bestimmt (**Abb. 3-12**). Dazu wird der Punkt maximaler Krümmung bzw. der Zyklenwert des Maximums der zweiten Ableitung des *Splines*, $c_{S''max}$, berechnet (**Abb. 3-12 A**). Der Datenpunkt 4 Zyklen vor diesem Punkt (c_A) ist der ge-

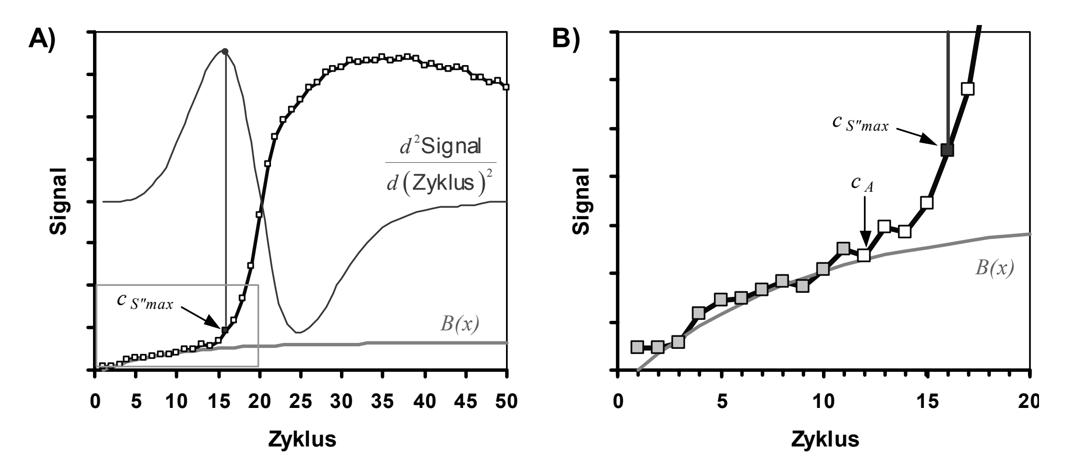


Abb. 3-12 Hintergrundkorrektur und Bestimmung der exponentiellen Phasen von Amplifikationskurven. Das Verfahren ist im Text beschrieben. **A)** Unkorrigierte Amplifikationsdaten (schwarz) und ihre zweite Ableitung (grau, dünne Linie) sowie der extrapolierte Hintergrund-Signaltrend $B(x)$ (grau, dicke Linie). **B)** Ausschnittsvergrößerung von A) (graues Rechteck). Die zur Interpolation von $B(x)$ verwendeten Datenpunkte sind hellgrau gefüllt.

schätzte Zykluswert des Beginns der exponentiellen Signalzunahme (**Abb. 3-12 B**). Bei einer Amplifikationseffizienz von 1.8 ist der Anteil des amplifikationsbedingten Signals beim Zykluswert $c_A - 1$ um mehr als eine Größenordnung geringer als beim Zykluswert $c_{S''max}$.

Die Parameter a, b und c einer Sättigungsfunktion der Form

$$B(x) = a \cdot \left(1 - e^{-bx}\right) + c \qquad \textbf{Gl. 3.10}$$

werden an die Datenpunkte vor c_A angepasst, um den nicht amplifikationsbedingten Signaltrend zu beschreiben (hellblau gefüllte Punkte in **Abb. 3-12 B**). Die Subtraktion der Funktionswerte für $B(x)$ von den Signalwerten resultiert in der korrigierten Amplifikationskurve, die allerdings vor dem Zykluswert c_A noch negative Werte aufweist, was nicht sinnvoll ist, da es keine negativen Fluoreszenzwerte geben kann. Daher wird noch der mittlere Wert des Signalrauschens, y_{noise}, zu jedem Datenwert addiert, um die korrigierten Amplifikationskurve zu erhalten.

Das durch y_{noise} beschriebene Signalrauschen wird anhand der Datenwerte bis zum Zyklus c_A der um die Hintergrundfunktion $B(x)$ korrigierten Amplifikationskurve abgeschätzt. Die Abweichungen der unkorrigierten Datenwerte von der angepassten Hintergrundfunktion für Werte kleiner dem Zykluswert c_A sind normalverteilt, wie am Beispiel einiger Signalkurven für Messungen mit Hybridisierungssonden gezeigt ist (**Abb. 3-13**).

Die Mittelwerte der berechneten Verteilungen liegen sehr nahe Null. Das bedeutet, dass die Funktion die amplifikationsunabhängigen Trends in den Signalkurven gut beschreibt, unabhängig davon, ob die Daten von Kanal 2 oder vom Verhältnis Kanal 2 / Kanal 1 verwendet werden. Eine systematische Abweichung der Funktion $B(x)$ würde in einer schiefen Normalverteilung oder in einer Verteilung der Differenzwerte resultieren, woraus sich das

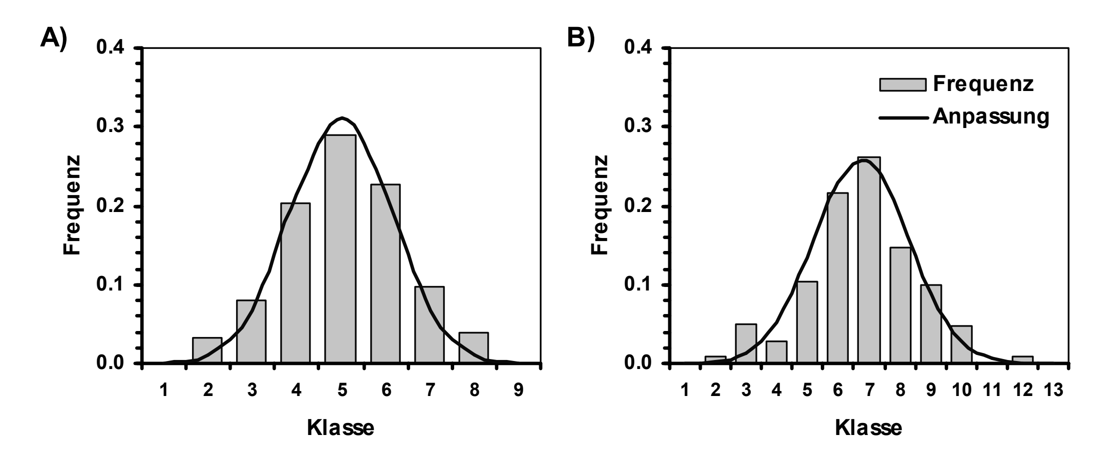


Abb. 3-13 Verteilungen der Differenzen der Datenwerte der Signalkurven von den Werten der angepassten Hintergrundfunktion in den Zyklen vor Beginn der erkennbaren exponentiellen Phase. Humane genomische DNA wurde mit dem Primerpaar IGF-H1/R1 amplifiziert. Das PCR-Produkt wurde mit den Hybridisierungssonden IGF-D/A detektiert. Die Differenzwerte $y_i – B(x_i)$ von 10 Amplifikationskurven vor Beginn der erkennbaren exponentiellen Phasen wurden ausgewertet ($1 \leq i \leq c_A$). Die Gesamtzahl der ausgewerteten Datenpunkte beträgt 133. Die dunkelblaue Linie zeigt die an die Histogramme angepasste Normalverteilung. **A)** Auswertung der Daten von Kanal 2/Kanal 1. Wertebereich: $-4 \cdot 10^{-4} \ldots 4 \cdot 10^{-4}$; Klassengröße: $0.73 \cdot 10^{-4}$. Standardabweichung der Normalverteilung: $2.02 \cdot 10^{-4}$. **B)** Auswertung der Daten von Kanal 2. Wertebereich: $-3 \cdot 10^{-3} \ldots 3 \cdot 10^{-3}$; Klassengröße: $0.4 \cdot 10^{-3}$. Standardabweichung der Normalverteilung: $1.13 \cdot 10^{-3}$.

Signalrauschen nicht ohne weiteres ableiten ließe. Mit y_{noise} als der doppelten Standardabweichung SD dieser Verteilung sind nach der Addition von y_{noise} zu den Datenwerten durchschnittlich 95 % der Datenwerte mit $i \leq c_A$ positiv.

Ein Beispiel für den Effekt der Korrektur auf die Signalkurven ist in **Abb. 3-14** dargestellt: Die korrigierten Signalkurven weisen längere exponentielle Phasen auf. Während die beobachteten Effizienzen bzw. die Steigungen der exponentiellen Abschnitte in der logarithmischen Auftragung bei den unkorrigierten Kurven mit abnehmender initialer Templatmenge erkennbar abnehmen bzw. flacher werden, verlaufen die log-linearen Abschnitte der Kurven nach Korrektur nahezu parallel. Dieser Effekt wird in **3.2.1.6** näher untersucht.

3.2.1.3 Bestimmung der exponentiellen Phase

Die Bestimmung der Phase erkennbarer exponentieller Signalzunahme erfolgt unter Verwendung der logarithmierten Werte der korrigierten Signalkurve (**Abb. 3-15**).

Die globalen Extrema bei den Zykluswerten $c_{e,min}$ und $c_{e,max}$ der zweiten Ableitung der logarithmierten Kurve schließen die erkennbare log-lineare bzw. exponentielle Phase ein.

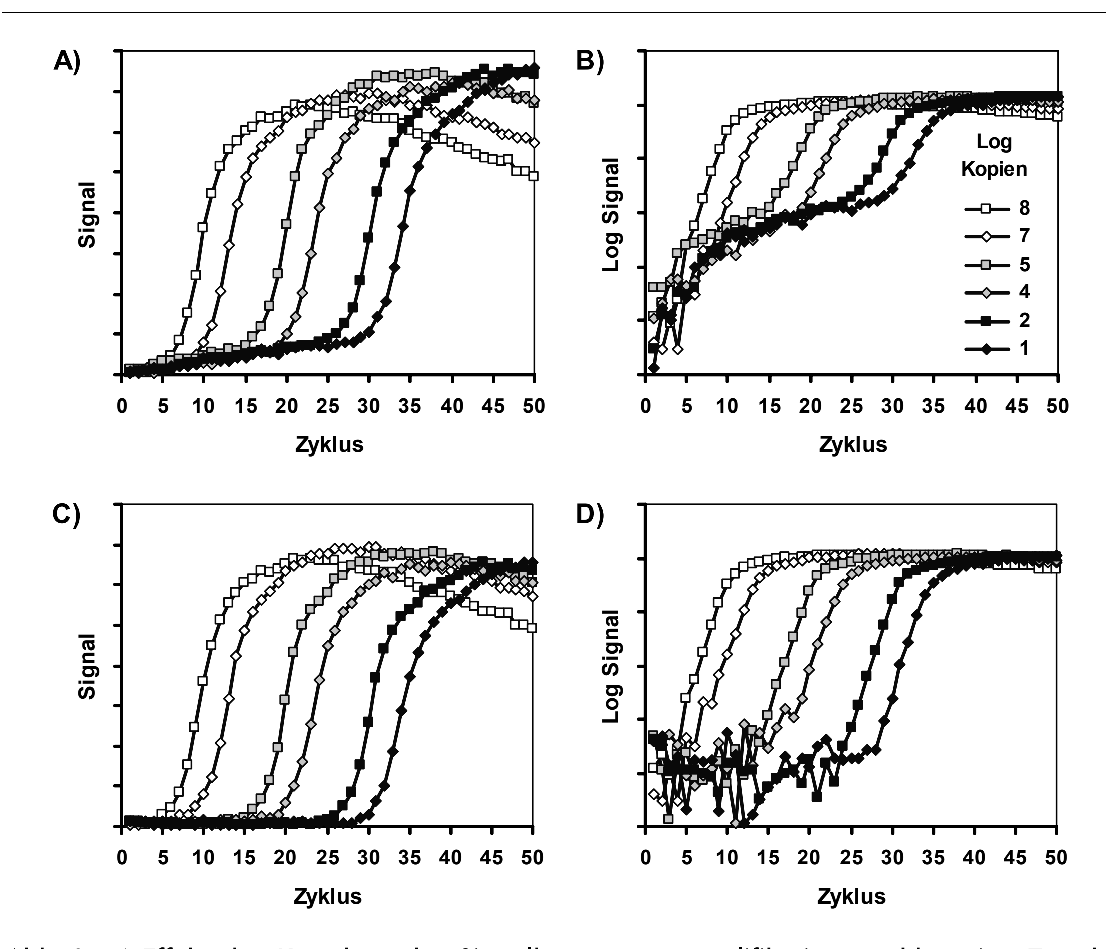


Abb. 3-14 Effekt der Korrektur der Signalkurven um amplifikationsunabhängige Trends. Die Signalkurven stammen von der Amplifikation der *IGF-1*-Zielsequenz mit den Primern IGF-H1/R1, ausgehend von aufgereinigtem PCR-Produkt als Templat-DNA. Die Detektion erfolgte mit *IGF-1*-spezifischen Hybridisierungssonden. **A)** Signalkurven von Kanal 2 / Kanal 1. **B)** Logarithmische Darstellung der Kurven aus A). **C)** Kurven aus A) nach Korrektur der amplifikationsunabhängigen Signaltrends. **D)** Logarithmische Darstellung der Kurven aus C).

Zur Berechnung der zweiten Ableitung werden die logarithmierten Daten, wie in **3.2.1.2** beschrieben, durch ein glättendes *Spline* interpoliert. Innerhalb dieses Bereiches werden nur diejenigen Punkte zur erkennbaren exponentiellen Signalzunahme gezählt, die sich oberhalb $2 \cdot y_{noise}$ befinden. Die Wahrscheinlichkeit, dass ein Datenpunkt zufällig, d. h. rauschbedingt, über dieser Grenze liegt, beträgt etwa 5 %, da die trendkorrigierten Signalwerte zunächst um $y_{noise} = 2 \cdot SD$ verschoben wurden.

3.2.1.4 Bestimmung des C_T-Wertes

Zur Berechnung des C_T-Wertes einer Probe wird eine Amplifikationsfunktion $A(x)$ an die korrigierten Datenwerte innerhalb der erkennbaren exponentiellen Phase angepasst. Bei der Amplifikationsfunktion kann es sich im einfachsten Falle um die Exponentialfunktion (**Gl. 1.4**) handeln, die Übergänge zu den angrenzenden Datenwerten können mit sigmoi-

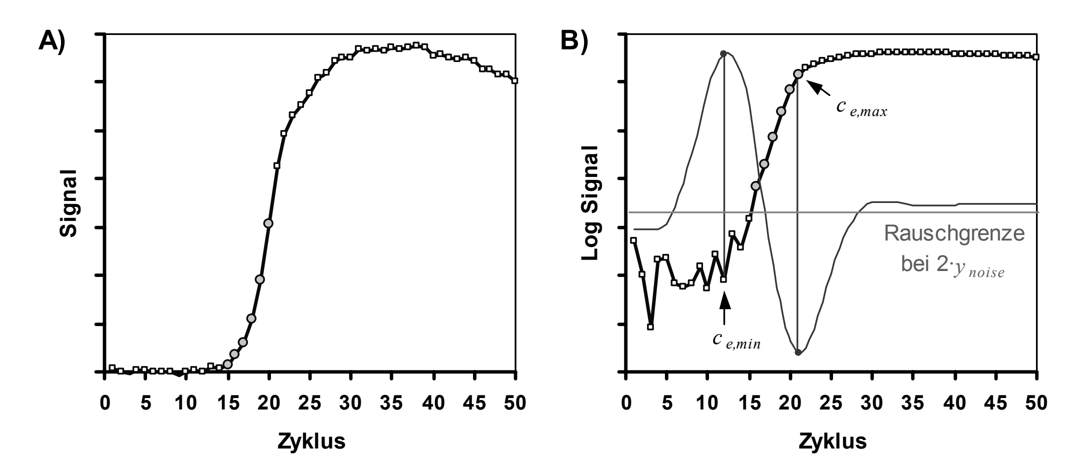


Abb. 3-15 Bestimmung der Phase erkennbarer exponentieller Signalzunahme. A) Eine um den Hintergrund-Signaltrend korrigierte Amplifikationskurve. **B)** Logarithmische Auftragung der Signalkurve aus A) (schwarz) und die zweite Ableitung der logarithmierten Daten (grau) sowie die Rauschgrenze (horizontale Linie). Die Datenpunkte innerhalb der bestimmten exponentiellen Phase sind hellgrau gefüllt.

den Funktionen angepasst werden. Der C_T-Wert ist definiert als der Zykluswert, bei welchem die Signalkurve einen Schwellenwert S überschreitet. Er kann auf einfache Weise bestimmt werden, indem die Gleichung $A(C_T) = S$ nach C_T aufgelöst wird. Der Wertebereich von S beschränkt sich dabei sinnvollerweise auf den Wertebereich der erkennbaren exponentiellen Phase. Für größere oder kleinere Werte von S kann die Amplifikationsfunktion $A(x)$ signifikant von dem Verlauf der Signalkurve abweichen.

3.2.1.5 Bestimmung des optimalen Schwellenwertes

Wie bereits in **3.1.2** und **3.1.3** gezeigt, wirken sich bei niedrigen Schwellenwerten vorwiegend das Signalrauschen, bei hohen Schwellenwerten vorwiegend die Plateauphasen negativ auf die Genauigkeit der Quantifizierung und mithin auf den Fehler der Kalibriergeraden aus. Zwischen den beiderseits extremen Einflüssen muss ein Schwellenwert existieren, der eine Kalibriergerade mit minimalem Fehler liefert.

Anhand der C_T-Werte der Standards mit bekannter initialer Templatmenge wird wie in **0** beschrieben eine Kalibriergerade ermittelt. Der Definitionsbereich des Schwellenwerts ist der Überlappungsbereich der exponentiellen Phasen aller auszuwertenden Standards. Für jeden beliebigen Schwellenwert innerhalb dieses Definitionsbereiches kann eine Fehlergröße der Kalibriergeraden, z. B. die Fehlerquadratsumme, angegeben werden. Da C_T-Werte in Abhängigkeit von der Lage des Schwellenwertes für alle Signalkurven eine wertekontinuierliche Funktion beschreiben, ist auch die Fehlerfunktion der Kalibriergeraden wertekontinuierlich. Sie besitzt genau ein Minimum im Definitionsbereich. Dies ist ein entscheidender Vorteil gegenüber der herkömmlichen *Fit-points*-Methode (siehe **3.2.1.1**), die keine kontinuierliche Fehlerfunktion der Standardgeraden liefert, da je nach Lage des Schwellenwertes andere Datenpunkte der Signalkurven zur Berechnung der C_T-Werte verwendet werden.

Beispiele für die Fehlerfunktionen von Standardkurven sind in **Abb. 3-16** gezeigt. Sowohl für Daten von SYBR-Green I-Messungen (**Abb. 3-16 A**) als auch von Messungen mit Hybridisierungssonden (**Abb. 3-16 B**) sind die Wertebereiche der Fehlerfunktionen für beide C_T-Berechnungsmethoden (über die Amplifikationsfunktion $A(x)$ oder nach der *Fit-points*-Methode) vergleichbar. Man erkennt bei den in **Abb. 3-16** dargestellten Fehlerfunktionen nach Auswertung mit der *Fit-points*-Methode unterschiedlich breite kontinuierlicher Abschnitte mit dazwischen liegenden Diskontinuitätsstellen (Sprüngen). Diese Schwellenwerte der Sprünge liegen genau auf den Datenpunkten der Signalkurven.

Im Vergleich weisen die Standardgeraden von SYBR-Green I-Messungen meistens geringere Fehlerwerte auf als die von Messungen mit Hybridisierungssonden. Das ist möglicherweise auch durch die geringeren erreichten absoluten Signalwerte von Hybridisierungssonden begründet, was ein geringeres Signal/Rausch-Verhältnis bedingt (s. a. **3.1.2**). Mit SYBR-Green I lassen sich generell höhere Signalintensitäten erreichen als mit sequenzspezifischen Sonden, da hier je Amplikon im Reaktionsansatz eine Vielzahl von SYBR-Green I-Molekülen mit hoher Quantenausbeute zur Fluoreszenz beitragen, wohingegen bei sequenzspezifischen Sonden je Amplikon nur etwa ein Sondenmolekül (beim TaqMan-Format zwei: die Summe aller in den vorherigen Zyklen hydrolysierten Sonden ist bei einer Effizienz von zwei gerade gleich der Menge der im aktuellen Zyklus hydrolysierten) fluoresziert.

Ein weiteres mögliches Kriterium für den optimalen Schwellenwert ist die maximale Präzision (bzw. die minimale Streuung) der C_T-Werte identischer Ansätze. Wie für den Fehler der

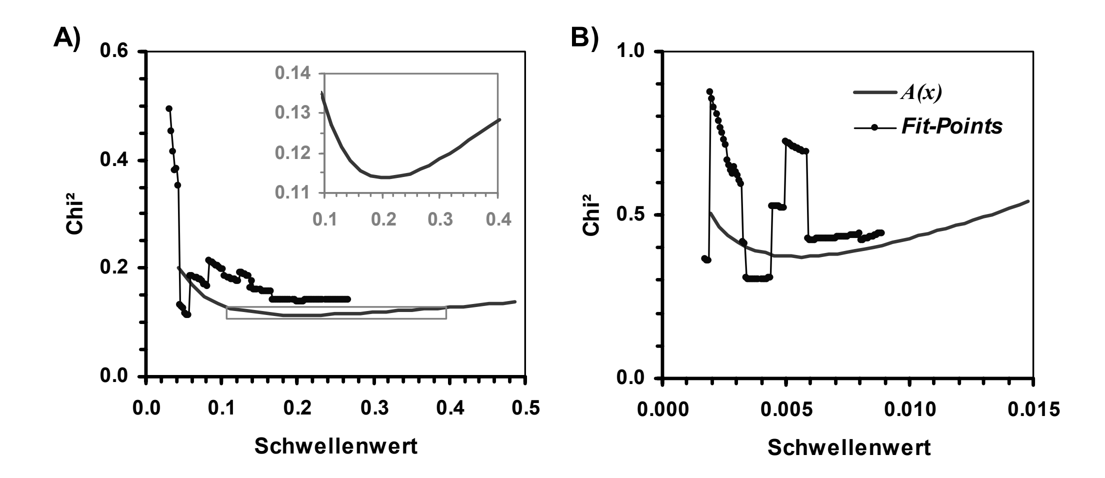


Abb. 3-16 Fehlerfunktion der Standardgeraden in Abhängigkeit des Schwellenwertes. Die C_T-Werte wurden von den korrigierten Signalkurven der Standardreihen berechnet. Die Berechnung erfolgte unter Verwendung der an die exponentiellen Phasen angepassten Exponentialfunktion $A(x) = a \cdot \varepsilon^x$ bzw. nach der *Fit-points*-Methode. Die Signalkurven wurden den Demo-Dateien aus dem Lieferumfang des LightCyclers entnommen. **A)** SYBR-Green I-Messung (Signalwerte von Kanal 1): „Demo-Quant SYBR-Green“. Das Minimum der kontinuierlichen Fehlerfunktion ist zur Verdeutlichung in der Ausschnittsvergrößerung dargestellt. **B)** Hybridisierungssonden (Signalwerte: Kanal 2 / Kanal 1): „Demo-Quantification-RED-640“.

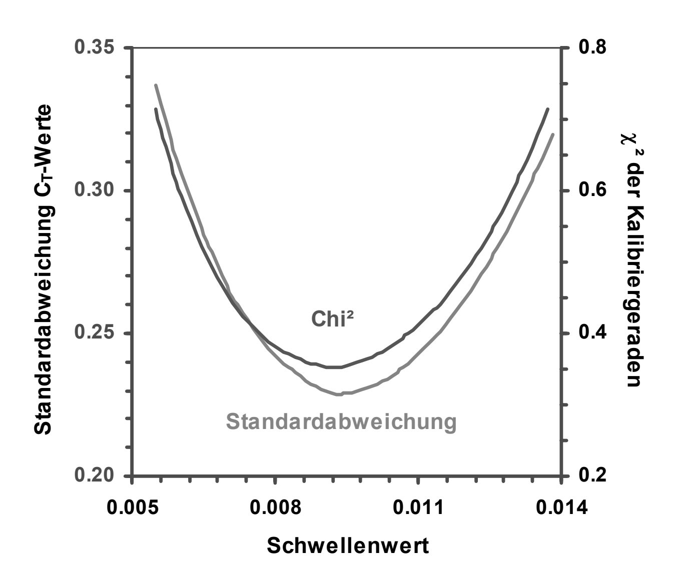


Abb. 3-17 Präzision der C_T-Werte und Fehler der Kalibriergeraden in Abhängigkeit vom Schwellenwert. Die Standardabweichung der C_T-Werte von 5-fach-Ansätzen dient als Maß für die Präzision, die Fehlerquadratsumme χ^2 der Kalibriergeraden als Maß für den Fehler der Kalibriergeraden. IGF-1-Zielsequenz nach Amplifikation mit den Primern IGF-H1/R1 und Detektion mit dem Sondenpaar IGF-D/A. Ausgewertet wurden die Signalwerte von Kanal 2/Kanal 1.

Kalibriergeraden existiert auch für die Streuung der C_T-Werten von Mehrfachansätzen ein optimaler Schwellenwert mit minimaler Streuung. Für Experimente, die sowohl Standards als auch Mehrfachansätze von Proben enthalten, zeigt sich, dass die Lagen des Schwellenwerts bei einem minimalen Fehler der Standardgeraden und einer minimalen Streuung der C_T-Werte von Mehrfachansätzen etwa gleich sind, wie in **Abb. 3-17** beispielhaft für eine Messung gezeigt ist.

3.2.1.6 Vergleichende Auswertungen mit dem neuen Algorithmus

In diesem Abschnitt werden die Ergebnisse von Untersuchungen zur Genauigkeit und Robustheit des entwickelten Algorithmus dargestellt und mit den Ergebnissen der Auswertung nach der *Fit-points*-Methode verglichen.

Wie schon in **Abb. 3-14** zu sehen, bewirkt die Korrektur der Signalkurven eine Verlängerung der erkennbaren exponentiellen Abschnitte, was die Verwendung von mehr Punkten zur Berechnung der C_T-Werte ermöglicht und dadurch die Rauschfilterung verbessert. Weiterhin lässt sich erkennen, dass die log-linearen Abschnitte der unkorrigierten Signalkurven nicht parallel verlaufen. Die Steigung verringert sich offensichtlich mit abnehmender initialer Templatmenge. Die Steigung der Log-linearen Abschnitte entspricht der gemessenen Amplifikationseffizienz der individuellen Reaktion. Diese müssen aber in der quantitativen RT-PCR für alle initialen Templatmengen gleich sein, damit der lineare Zusammenhang zwischen dem C_T-Wert und dem Logarithmus der initialen Templatmenge gegeben ist.

Eine genaue Untersuchung der Abhängigkeit der aus den individuellen Signalkurven berechneten Effizienzen zeigt bei den unkorrigierten Kurven eine deutliche Abnahme mit der initialen Templatmenge, während die von den korrigierten Signalkurven berechneten Effizienzen nicht mit der Templatmenge korrelieren (**Abb. 3-18 A**). Wenn die log-linearen Abschnitte wie bei den unkorrigierten Signalkurven nicht parallel sind, ist die aus der Steigung der Kalibriergeraden (**Gl. 1.7**, $m = -(\log(\varepsilon))^{-1}$) berechnete Effizienz abhängig von der Höhe des Schwellenwertes (**Abb. 3-18 B**). Für die korrigierten Signalkurven ist die so berechnete Effizienz nahezu unabhängig von der Lage des Schwellenwertes und stimmt sehr gut mit den aus den individuellen Signalkurven ermittelten Effizienzen überein.

Die erreichbare Präzision in der Berechnung der C_T-Werte hängt von der verwendeten Amplifikationsfunktion ab, mit welcher die exponentiellen Abschnitte der Signalkurven beschrieben werden. So können z. B. sigmoide Funktionen die Randbereiche der exponentiellen Abschnitte besser berücksichtigen. Abweichungen vom exponentiellen Signalverlauf in diesen Randbereichen haben auf sigmoide Funktionen einen eher lokalen Einfluss, während sie die Lage der einfachen Exponentialfunktion global stärker beeinflussen. Die erreichbare Präzision wurde für die folgenden Amplifikationsfunktionen ermittelt:

$$A(x) = N_0 \cdot e^{r \cdot x} \qquad \textbf{Gl. 3.11}$$

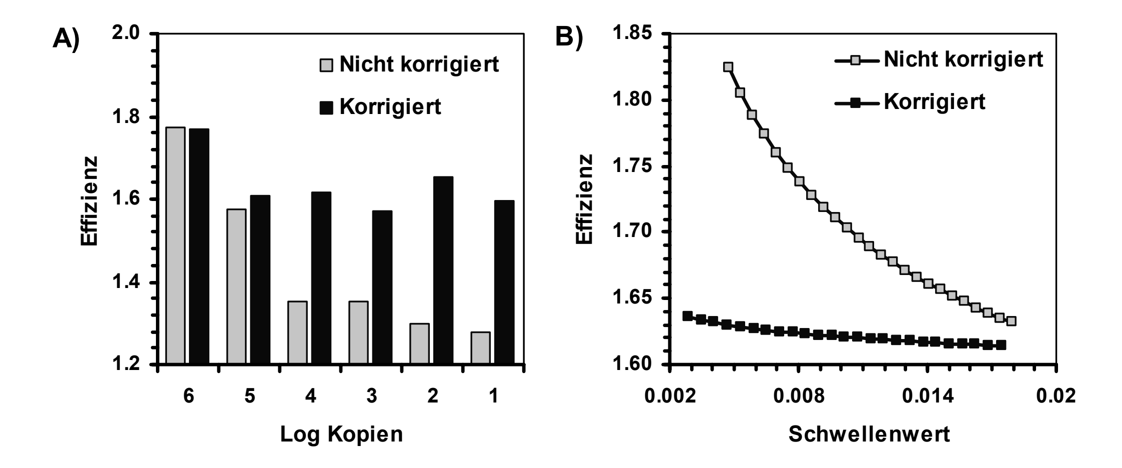


Abb. 3-18 Berechnete Amplifikationseffizienzen in Abhängigkeit der initialen Templatmenge. Amplifiziert wurde die *IGF-1*-Zielsequenz mit den Primern IGF-H1/R1 ausgehend von PCR-Produkt als initialem Templat, die Detektion erfolgte durch die *IGF-1*-spezifischen Hybridisierungssonden (Signalwerte: Kanal 2 / Kanal 1). **A)** Die Effizienzen wurden anhand der individuellen Signalkurven ermittelt. Dazu wurde die Exponentialfunktion $A(x) = a \cdot \varepsilon^x$ an die wie in 3.2.1.3 bestimmte exponentielle Phase jeder Kurve angepasst. Der Parameter ε entspricht der Amplifikationseffizienz. MW und SD der korrigierten Daten sind 1.64 ± 0.07. **B)** Amplifikationseffizienzen in Abhängigkeit des Schwellenwerts. Die Effizienzen wurden anhand der Steigungen ($m = -(\log(\varepsilon))^{-1}$) der Kalibriergeraden (Gl. 1.7) für verschiedene Schwellenwerte berechnet.

$$A(x) = N_{\max} \cdot \frac{x^r}{x^r + x_T{}^r} \qquad \text{Gl. 3.12}$$

$$A(x) = N_{\max} \cdot \frac{1}{1 + e^{-r \cdot (x_T - x)}} \qquad \text{Gl. 3.13}$$

Die Variable x bezeichnet die Zykluszahl, N_0 die initiale Templatmenge, N_{max} die PCR-Produktmenge in der Plateauphase. Der Parameter r ist ein Maß für die Amplifikationseffizienz, x_T ist die Zykluszahl beim Wendepunkt der Funktion. Der Parameter r entspricht nicht exakt der Amplifikationseffizienz ε. In **Gl. 3.11** ist $r = \ln(\varepsilon)$, während r in **Gl. 3.12** und **Gl. 3.13** die Steigung der Funktion im Wendepunkt ($x = x_T$) anpasst. Der Vergleich der drei angegebenen Amplifikationsfunktionen hinsichtlich der erreichbaren Präzision in der Berechnung der C_T-Werte wurde an einer aus 32 Signalkurven durch Medianbildung gemittelten Kurve durchgeführt, der ein Signalrauschen überlagert wurde.

Durch die Medianbildung wird das reale Signalrauschen verringert, ohne die Charakteristik des Kurvenverlaufs zu verändern. Die Signalkurven zur Medianbildung stammen von einer Amplifikation der p53-Zielsequenz ausgehend von 30 ng humaner genomischer DNA als

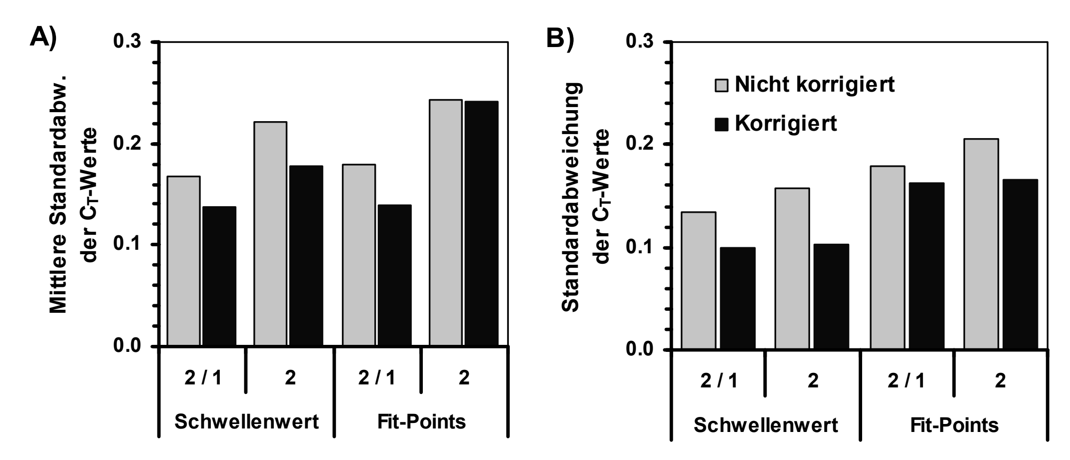


Abb. 3-19 Präzision der C_T-Werte. Amplifiziert wurde die *p53*-Zielsequenz aus humaner genomischer DNA, die Detektion erfolgte mit Hybridisierungssonden. Die C_T-Werte wurden entweder anhand der Signalkurven aus Kanal 2 oder dem Signalverhältnis von Kanal 2 / Kanal 1 nach der in 3.2.1.4 beschriebenen Schwellenwertmethode (angepasste Amplifikationsfunktion: Gl. 3.13) oder der *Fit-points*-Methode (siehe 3.2.1.1) mit 3 Datenpunkten je Kurve berechnet. **A)** Die Berechnung der C_T-Werte erfolgte jeweils mit dem Schwellenwert bzw. dem *Noiseband*, für den der Fehler der Kalibriergeraden minimal war (siehe 3.2.1.5). Gezeigt sind SD der C_T-Werte von Fünffach-Ansätzen, gemittelt von fünf Proben mit unterschiedlichen initialen Templatmengen (zwischen 15 und 30 ng humane genomische DNA) und aus fünf unabhängigen Experimenten (n = 125). **B)** Standardabweichung der C_T-Werte von 32 identischen Ansätzen mit 1000 Templatkopien je Reaktion (Die Experimente enthielten keine Standards). Der Schwellenwert wurde dabei in die Mitte und der *Noiseband* an die Untergrenze des Überlappungsbereichs aller exponentiellen Phasen gelegt.

Templat. Die Detektion erfolgte durch Hybridisierungssonden, verwendet wurde das Signalverhältnis Kanal 2 / Kanal 1. Zur Bestimmung der Präzision wurden die SD der CT-Werte von jeweils 50 Mediankurven mit überlagertem Rauschen unterschiedlicher Intensität berechnet. Als Rauschanteile dienten Zufallswerte, die nach der Box-Muller-Methode [210] berechnet wurden. Der Schwellenwert wurde in die Mitte der exponentiellen Phase gesetzt. Für Rauschintensitäten mit SD von bis zu 0.5 % der Plateausignalwerte lieferte die einfache Exponentialfunktion (**Gl. 3.11**) 10–15 % präzisere Ergebnisse als die sigmoide logistische Funktion (**Gl. 3.12**). Bei größeren Rauschintensitäten war die logistische Funktion präziser. In allen Fällen war jedoch die Präzision der C_T-Werte bei Verwendung der sigmoiden exponentiellen Funktion (**Gl. 3.13**) 10-25 % größer als bei Verwendung der anderen getesteten Funktionen.

Die Präzision der ermittelten C_T-Werte erhöht sich auch durch die Korrektur der amplifikationsunabhängigen Signaltrends, wie es sich aus der SD der C_T-Werte von Mehrfachansätzen ableiten lässt (**Abb. 3-19**). In allen Fällen ist die Präzision der C_T-Werte bei der Bestimmung nach der Schwellenwertmethode mindestens ebenso hoch oder höher als bei Verwendung der Fit-points-Methode. Im allgemeinen werden bei Detektion mit Hybridisierungssonden höhere Präzisionen erreicht, wenn die Daten des Verhältnisses Kanal 2 / Kanal 1 anstelle der Daten von Kanal 2 verwendet werden. Bei der Verhältnisbildung kürzen sich diejenigen Einflüsse heraus, die sich gleichermaßen auf das Donorsignal (gemessen in Kanal 1) wie auf das Akzeptorsignal (gemessen in Kanal 2) auswirken. Dazu zählen insbesondere Intensitätsschwankungen durch Abweichungen bei den Positionierungen der Kapillaren über dem Linsensystem des LightCycler-Instruments.

Die Robustheit der Auswertung bei verrauschten Signalkurven wurde wie oben beschrieben an einer durch Medianbildung von Signalkurven gewonnenen Kurve mit überlagertem Rauschen unterschiedlicher Intensität gewonnen. Die SD des Rauschens wurde zwischen 0.17 % und 1.36 % der Signalwerte der Mediankurve variiert. Dabei zeigte sich, dass die Präzision der ermittelten C_T-Werte wie erwartet mit zunehmender Rauschintensität bei beiden Verfahren (Schwellenwertmethode und *Fit-points*-Methode) abnimmt und die Schwellenwertmethode immer eine ca. 20 % größere Präzision erreicht als die *Fit-points*-Methode.

Der Gewinn an Genauigkeit und Präzision des hier entwickelten Algorithmus im Vergleich zum herkömmlichen Verfahren ist in **Abb. 3-20** demonstriert. Während beide Verfahren bei Auswertungen qualitativ sehr guter Messungen sehr vergleichbare Ergebnisse liefern, erreicht das hier entwickelte Verfahren bei Messungen mit geringeren Signal/Rausch-Verhältnissen eine deutlich höhere Genauigkeit bei gleichzeitig größerer Präzision.

Im gezeigten Beispiel erhält man nach dem herkömmlichen Verfahren für eine Probe mit real 800 Templatkopien einen MW von 544 Kopien für die Test-Probe, bei einer SD von 379 Kopien. Werden dieselben Daten nach dem hier entwickelten Verfahren ausgewertet, beträgt das Ergebnis 813 $\pm$ 115 Kopien (MW $\pm$ SD).

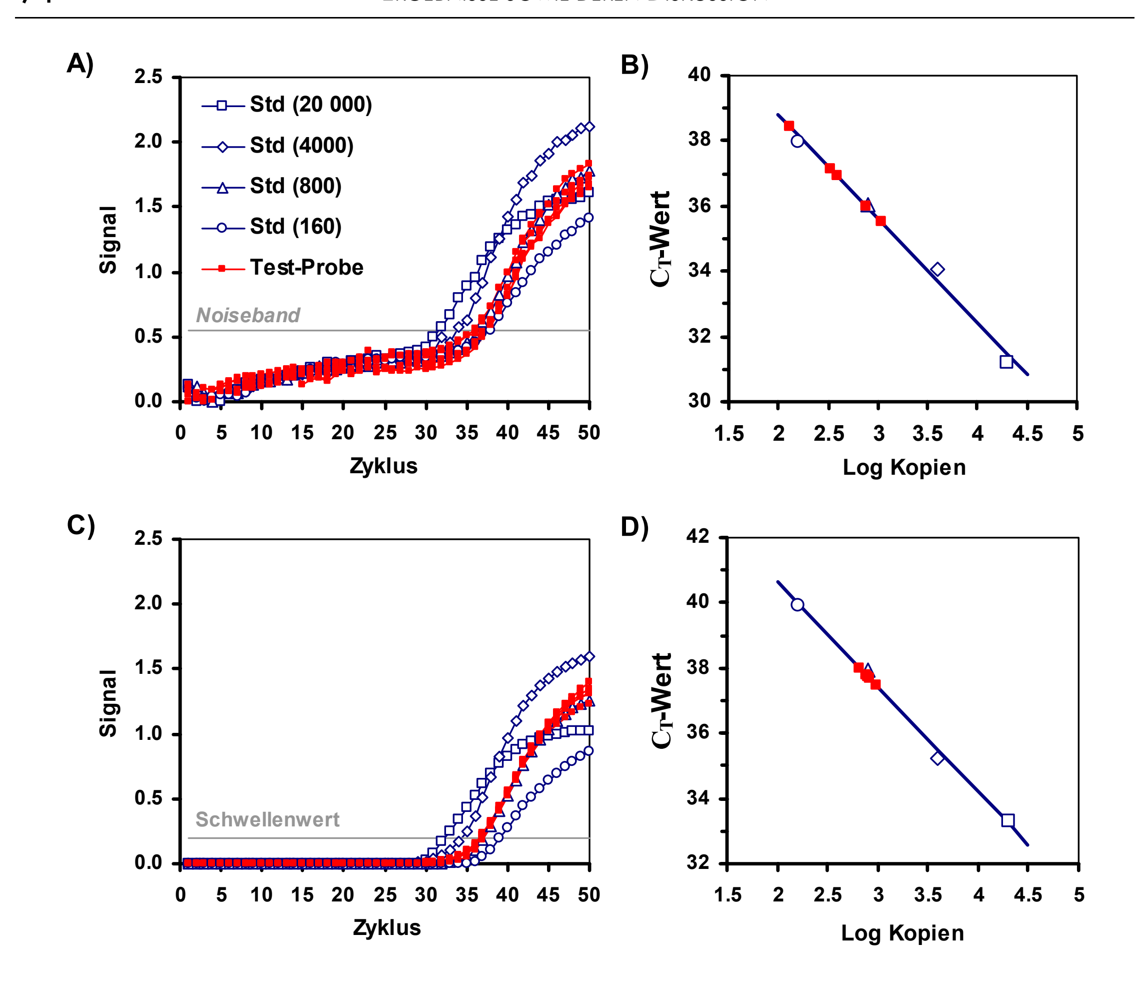


Abb. 3-20 **Genauigkeit und Präzision von Quantifizierungen.** Die *p53*-Zielsequenz wurde aus humaner genomischer DNA amplifiziert. Die Detektion erfolgte mit Hybridisierungssonden, zur Auswertung wurde das Verhältnis der Messwerte von Kanal 2 / Kanal 1 verwendet. **A)** Unkorrigierte Amplifikationskurven einer seriellen Verdünnungsreihe sowie von einem Fünffach-Ansatz einer Test-Probe. Bei der Test-Probe handelt es sich um die Standard-Probe mit 800 Kopien, die es zu bestimmen gilt. Die horizontale Linie gibt die Lage des Noiseband an. Jeweils drei Datenwerte jeder Kurve direkt oberhalb des *Noiseband* wurden für die nach der *Fit-points*-Methode berechneten C_T-Werte verwendet. **B)** Kalibriergerade anhand der nach der *Fit-points*-Methode berechneten C_T-Werte der Signalkurven aus A). Die roten Datenpunkte bezeichnen die auf der Kalibriergeraden abgebildeten C_T-Werte der Test-Proben. **C)** Signalkurven aus A) nach Korrektur amplifikationsunabhängiger Signaltrends. Die horizontale Linie zeigt die Lage des Schwellenwerts zur Berechnung der C_T-Werte nach Anpassung der Amplifikationsfunktion. **D)** Wie B), nur von den korrigierten Signalkurven. Die Berechnung der C_T-Werte erfolgte nach der Schwellenwertmethode mit der in Gl. 3.13 beschriebenen Amplifikationsfunktion.

3.2.2 Algorithmus zur Auswertung von Schmelzkurven

3.2.2.1 Prinzip

Schmelzkurven werden aufgezeichnet, indem während eines langsamen Temperaturanstiegs die Fluoreszenzintensität der Reaktionsansätze kontinuierlich gemessen wird. Dies geschieht zumeist im Anschluss an die Amplifikation einer zu analysierenden Zielsequenz. Je nach verwendetem Sondensystem wird bei der Schmelzkurvenanalyse entweder die Schmelztemperatur (T_m) des PCR-Produkts (mit SYBR-Green I) bzw. der Sonden/Templat-Hybride (bei Hybridisierungssonden und *Molecular Beacons*, nicht allerdings bei TaqMan-Sonden und *Sunrise*- bzw. *Scorpion-Primern*) bestimmt. Bei der Schmelztemperatur der Produkte bzw. Hybride zeigen die gemessenen Signalkurven ein lokales Maximum in der Abnahme der Intensität. Zur besseren Identifizierbarkeit der maximalen Intensitätsänderungen werden die Schmelzkurven als Peakkurven dargestellt. Bei den Peakkurven handelt es sich um die negative erste Ableitung der Schmelzkurven nach der Temperatur. Die einfach zu erkennende Lage der Peakmaxima entspricht dann der Schmelztemperatur des Produkts.

Unter der Voraussetzung, dass die gemessenen Signaländerungen ausschließlich auf dem Schmelzvorgang beruhen, ist die Größe der Signaländerung über den Temperaturbereich eines Peaks proportional zur Produkt- bzw. Hybridmenge mit einem T_m-Wert innerhalb dieses Temperaturbereichs. Diese Signaländerung entspricht dem negativen Integral der Peakkurve in den Grenzen des Temperaturbereichs bzw. der Fläche des betreffenden Peaks.

Tatsächlich führt allerdings schon die bloße Temperaturzunahme, unabhängig vom eigentlichen Schmelzvorgang, zu einer Abnahme der Fluoreszenzintensität. Durch die mit der Temperatur zunehmende Molekülbewegung vergrößert sich die Frequenz intermolekularer Stöße sowie deren mittlere Energie, wodurch zunehmend mehr Anregungsenergie der Fluorophore durch Stoßprozesse strahlungsfrei abgegeben wird.[216] Die Größe dieses temperaturabhängigen Quenchs ist abhängig von den Pufferbedingungen und von der Menge fluoreszierender Moleküle. Dieser Quench hat damit den Effekt, dass die Menge eines Produkts/Hybrids mit höherem T_m in einem Gemisch zu niedrig eingeschätzt wird, was bei geringen Mengen dieses Produkts im Vergleich zur Menge der Produkte mit niedrigeren T_m-Werten sogar dazu führen kann, dass es sich in der Schmelzkurvenanalyse nicht nachweisen lässt.

3.2.2.2 Korrektur der temperaturabhängigen Signaländerung

Die Messdaten werden zunächst durch glättende, kubische nichtparametrische *Splines* beschrieben[214,215] (vgl. **3.2.1.2**). Die Aufgabe der *Spline*-Funktion besteht hier in erster Linie in der Glättung der Originaldaten, da sich das Signalrauschen empfindlich auf die Ableitung auswirkt. Der Glättungsparameter zur Berechnung der Ableitung wird in Abhängigkeit von der Anzahl zur Verfügung stehender Datenpunkte pro Messkurve (n) gewählt und beträgt $1000 \cdot n^{-1}$.

Der Intensitätsabfall des Schmelzkurvensignals kann betrachtet werden als die Summe zweier Anteile: einem Anteil proportional zum Schmelzprozess der fluoreszierenden Produkte und einem Anteil proportional zum temperaturbedingten Quench, der seinerseits wiederum abhängt von der Menge noch fluoreszierender Produkte im Reaktionsansatz. Um diese beiden Anteile voneinander differenzieren zu können, werden zunächst die ungefähren Charakteristika der Peaks in den Peakkurven bestimmt. Dabei werden die Peaks anhand lokaler Maxima in den Peakkurven identifiziert. Die angrenzenden Minima definieren die Grenzen des Peaks. Beträgt der Signalunterschied über den Temperaturbereich des gefundenen Peaks weniger als 10 % des Signals vor dem Peak, wird der Peak nicht berücksichtigt. In den Temperaturintervallen zwischen den Peaks ist die Menge fluoreszierender Produkte konstant, so dass hier alle Änderungen der Fluoreszenzintensität ausschließlich durch die Temperaturzunahme verursacht und nicht von Schmelzprozessen überlagert werden.

Innerhalb der Temperaturbereiche der gefundenen Peaks tragen sowohl der temperaturabhängige Quench als auch der Schmelzprozess zu der beobachteten Signalabnahme bei. Der temperaturabhängige Quench an den Grenzen dieser Temperaturbereiche ist bekannt (siehe oben). Der ungefähre Verlauf des Schmelzprozesses geht aus der aufgezeichneten Schmelzkurve innerhalb des Temperaturbereichs des Peaks hervor. Da die absolute Größe des temperaturabhängigen Quenchs proportional ist zur Menge der fluoreszierenden, also nicht geschmolzenen Produkte, entspricht der Quench im Peakintervall in guter Näherung der Abbildung der Schmelzkurve auf die Quenchwerte an den Intervallgrenzen.

Das Ergebnis einer Berechnung ist in **Abb. 3-21** veranschaulicht. Der ermittelte Quench lässt sich als Anteil der Peakkurvenwerte verstehen. Außerhalb der durch die Peakgrenzen

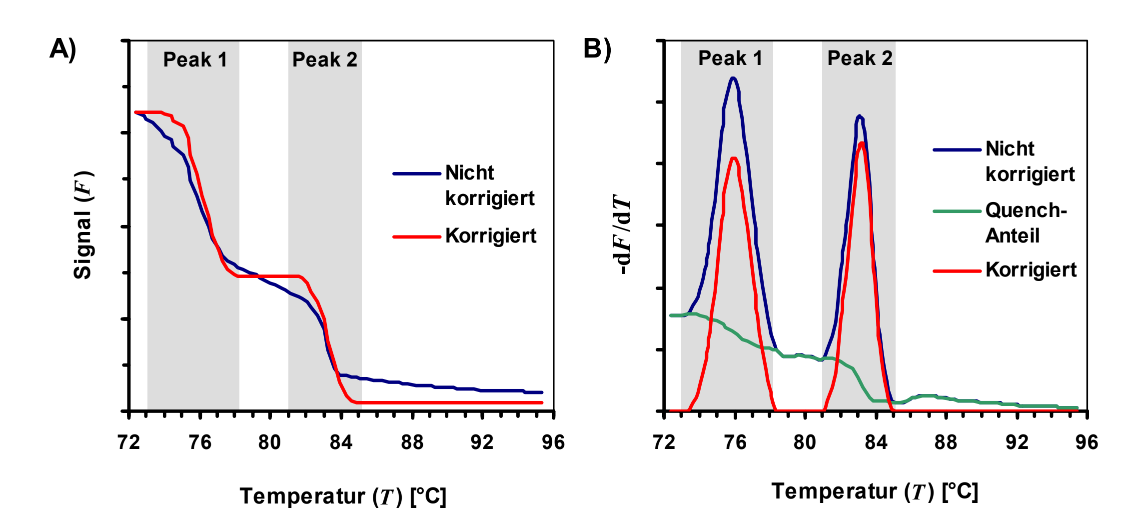


Abb. 3-21 Korrektur des temperaturabhängigen Quenchs von Schmelzkurvendaten. **A)** Schmelzkurven vor (blau) und nach (rot) Korrektur des temperaturabhängigen Quench. **B)** Peakkurven der Daten von A) sowie der ermittelte Quenchanteil (grün). Die Bereiche der Peaks sind hellgrau hinterlegt.

definierten Temperaturbereiche ist der Quench identisch mit der Peakkurve, innerhalb ist er proportional zur Schmelzkurve in diesem Temperaturbereich und verbindet praktisch die angrenzenden Segmente (grüne Kurve in **Abb. 3-21 B**).

Die Korrektur des Quenchs erfolgt durch die Subtraktion der berechneten Quenchwerte von der Peakkurve. Die korrigierte Peakkurve wird anschließend wieder integriert und auf den Signalbereich der ursprünglichen Schmelzkurve abgebildet.

3.2.2.3 Bestimmung der Menge und Schmelztemperatur der PCR-Produkte

Menge und Schmelztemperatur der entstandenen PCR-Produkte lassen sich anhand der Peakkurven bestimmen. Da die Peaks mehrerer im selben Reaktionsansatz enthaltenen Produkte sich überschneiden können, werden die Charakteristika der Peaks durch die Anpassung von **Gl. 3.14** an die korrigierte Peakkurve ermittelt:

$$y(T) = \sum_{i=1}^{n} h_i \cdot e^{-\left(\frac{T-\mu_i}{\sigma_i}\right)^2} \qquad \text{Gl. 3.14}$$

Die Funktion beschreibt eine Summe von Gaußkurven. Dabei ist T die Temperatur, n die Anzahl der Peaks, und h_i, μ_i und σ_i sind die Höhe, Schmelztemperatur und die Breite (SD) von Peak i.

Die zur Produktmenge proportionale Fläche des Peaks i ergibt sich aus **Gl. 3.14** durch

$$A_i(h_i, \sigma_i) = h_i \cdot \sigma_i \cdot \sqrt{\pi} \qquad \text{Gl. 3.15}$$

Es werden nur Peaks berücksichtigt, deren Höhe über der durch das Signalrauschen definierten Grenze liegt. Das Signalrauschen wird anhand der Abweichungsquadrate von Messwerten und *Spline*-Funktion ermittelt.

3.2.2.4 Vergleichende Auswertung mit dem neuen Algorithmus

Die rein temperaturbedingte, nicht durch Schmelzprozesse fluoreszierender Produkte hervorgerufene Signalabnahme kann bei geringen Produktmengen die schmelzbedingte Signalabnahme überschreiten. Je höher die Schmelztemperatur des nachzuweisenden Produkts, desto geringer ist der durch den Schmelzprozess bedingte Anteil an der Gesamtsignalabnahme. Die Korrektur des temperaturabhängigen Quenchs von Schmelzkurvendaten erhöht hier die Sensitivität, da die gesamte verbleibende Signalabnahme ausschließlich dem Schmelzprozess zugeordnet wird. Je nach Qualität der Ausgangsdaten lässt sich durch die Korrektur die Sensitivität im Vergleich zum herkömmlichen Verfahren um das vier- bis achtfache steigern. Ein Beispiel dafür ist in **Abb. 3-22** gezeigt.

Das spezifische PCR-Produkt (125 Bp) hat eine Schmelztemperatur von ca. 86 °C. Während die Probe mit 10 initialen Templatkopien in der nicht korrigierten Darstellung keinen erkennbaren Peak bei dieser Temperatur zeigt, ist nach der Korrektur deutlich zu erkennen, dass spezifisches Produkt gebildet wurde (vergleiche rote Kurve in **Abb. 3-22 B** bzw. **D**). Erst nach der Korrektur wird ein weiterer Peak bei etwa 79 °C sowohl bei der Leerprobe (0

Kopien) als auch bei der Probe mit 10 Kopien erkennbar. Bei den Produkten mit dieser Schmelztemperatur handelt es sich wahrscheinlich um Primerdimere. Durch PAGE-Analyse ließ sich im Fall der Probe mit 100 initialen Templatkopien nur das spezifische Produkt nachweisen.

Der Einfluss der Korrektur des temperaturabhängigen Quenchs bei Schmelzkurvenmessungen mit SYBR-Green I-Detektion auf die Genauigkeit der Produktmengenbestimmung anhand der Peakflächenintegrale wurde an den *PAI*- und *F7c*-Zielsequenzen getestet. Die aufgereinigten PCR-Produkte, generiert mit den Primerpaaren PAI-H1/R1 und F7c-H1/R1 wurden nach UV-spektrophotometrischer Konzentrationsbestimmung in den Verhältnissen 1:4, 1:2, 1:1, 2:1 und 4:1 gemischt und zur Schmelzkurvenanalyse in PCR-Reaktionsgemische gegeben. Die Aufzeichnung der Schmelzkurven erfolgte wie in **2.5.5** beschrieben. Das *PAI*-Produkt zeigt einen Schmelzpeak mit einem Maximum bei ca. 83.5 °C, während das *F7c*-Produkt bei einer Temperatur von ca. 89.5 °C schmilzt. Das Ergebnis ist in **Abb. 3-23** dargestellt.

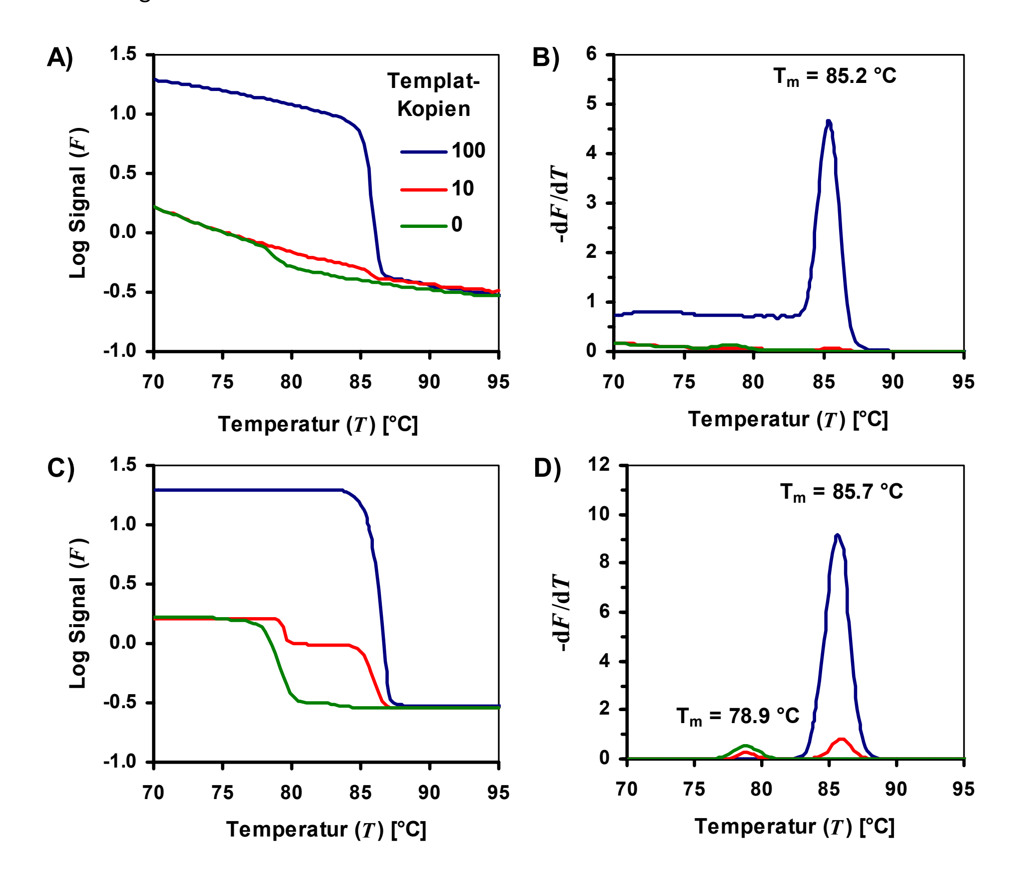


Abb. 3-22 **Steigerung der Sensitivität der Schmelzkurvenanalyse durch die Korrektur des temperaturabhängigen Quenchs.** Unterschiedliche Mengen humaner genomischer DNA wurden über 40 Zyklen mit den Primern IGF-H2/R2 in Anwesenheit von SYBR-Green I amplifiziert. Die PCR-Produkte wurden in einer anschließenden Schmelzkurvenanalyse untersucht. Die Detektion erfolgte in Kanal 1. **A)** Nichtkorrigierte Schmelzkurven. **B)** Peakkurven der Daten von A). **C)** Korrigierte Schmelzkurven. **D)** Peakkurven der Daten von C)

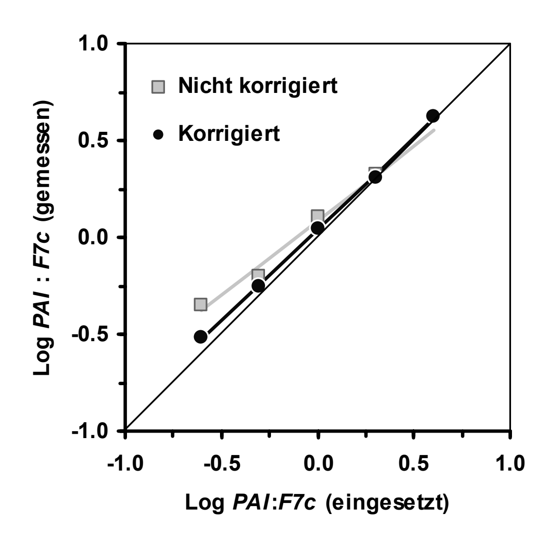


Abb. 3-23 Genauigkeit der Produktmengenbestimmung durch die Schmelzkurvenanalyse. Das Experiment ist im Text beschrieben. Aufgetragen sind die Verhältnisse der Flächenintegrale der Schmelzpeaks beider Produkte gegen das Verhältnis der eingesetzten Produktmengen sowie die zugehörigen Regressionsgeraden (nicht korrigiert: $y = 0.77 \cdot x + 0.08$; korrigiert: $y = 0.95 \cdot x + 0.04$).

Das im Unterschuss eingesetzte Produkt mit der niedrigeren Schmelztemperatur (*PAI*) wird ohne Korrektur des temperaturabhängigen Quenchs deutlich überbestimmt. Das liegt daran, dass bei seiner Schmelztemperatur das im Überschuss befindliche Produkt mit der höheren Schmelztemperatur (*F7c*) hier noch zur Gesamtfluoreszenz beiträgt und der dazu proportionale Quench die Signalminderung bei dieser Temperatur verstärkt. Das *PAI*-Produkt konnte ohne Korrektur bei einem vierfachen Überschuss des *F7c*-Produkts nicht mehr quantifiziert werden, da die Signalminderung bei der Schmelztemperatur (Peakhöhe) in diesem Fall geringer war als der anfängliche Signalverlust durch den Quench. Nach Korrektur des temperaturabhängigen Quenchs ist die tendenzielle Überbestimmung des Produkts mit der niedrigeren Schmelztemperatur nahezu eliminiert, die Steigung der Regressionsgeraden ist ungefähr gleich 1.

Die deutliche Steigerung der Genauigkeit durch die Quenchkorrektur bei der Bestimmung von Produktmengenverhältnissen anhand von Schmelzkurvenanalysen mit Hybridisierungssonden ist in **3.4.2.5** für den Fall einer kompetitiven PCR gezeigt. Dieses Beispiel demonstriert gleichzeitig die Vorzüge der Quenchkorrektur, nämlich die höhere Genauigkeit und den größeren dynamischen Bereich, für quantitative Schmelzkurvenanalysen, wie sie z. B. bei der quantitativen kompetitiven PCR verwendet werden.[147,217]

3.2.3 Algorithmus zur Farbkompensation

In RT-PCR-Experimenten können bei Verwendung sequenzspezifischer Sonden prinzipiell mehrere unterschiedliche PCR-Produkte im selben Reaktionsansatz nachgewiesen werden. Dazu müssen die jeweiligen spezifischen Sonden unterschiedliche Fluorophore tragen, deren Emissionslicht in separaten Kanälen detektiert wird. Sinnvoll ist dieses Vorgehen z. B. bei Multiplexansätzen in Hochdurchsatzverfahren, bei kompetitiven Amplifikationstechniken wie der quantitativen kompetitiven PCR oder der parallelen Bestimmung mehrerer Sequenzvarianten durch Schmelzkurvenanalysen mit *Molecular Beacons* oder Hybridisierungssonden (siehe auch **0** und **1.3**).

Ein Problem stellt hierbei die Bestimmung der Fluoreszenzintensitäten der individuellen Fluorophore aus den Messwerten der Kanäle dar. Die meisten verwendeten Fluorophore weisen partiell überlappende Emissionsspektren auf (**Abb. 2-3**, S. 39), so dass die in einem Kanal K_A gemessene Signalintensität nicht ausschließlich durch die Fluoreszenz desjenigen Fluorophors F_A bedingt ist, dessen Emissionsmaximum im Detektionsbereich des Kanals liegt, sondern möglicherweise noch die Emissionen anderer Fluorophore (F_B, F_C, usw.) zum gemessenen Signal beitragen. Die Größe dieses Beitrags z. B. von F_B zum Signal in K_A ist außer von der Fluoreszenzintensität von F_B auch von der Lage und Breite des durch die Filtersysteme definierten Detektionsbereichs des Kanals K_A sowie von der Form des Emissionsspektrums von F_B abhängig, welches wiederum von den Pufferbedingungen und der Temperatur abhängig ist.

Die Messung der Fluoreszenz eines Fluorophors in einem Kanal, dessen Detektionsbereich nicht das Emissionsmaximum des Fluorophors einschließt, wird als Übersprechen oder *crosstalk* bezeichnet. Die durch das Übersprechen bedingten Signalanteile können mit der Farbkompensation der Messwerte korrigiert werden, wenn das geräte-, puffer- und temperaturabhängige Übersprechverhalten der gleichzeitig verwendeten Fluorophore bekannt ist.

Das Übersprechverhalten wird in Farbkompensationsmessungen bestimmt, die nach einem für das jeweilige Gerät und die verwendeten Fluorophore vorgegebenen Protokoll durchgeführt werden. Ergebnis dieser Messung ist eine durch eine Kalibrierung bestimmte Parametermatrix, welche das temperaturabhängige Übersprechverhalten aller Fluorophore in allen Kanälen für ein bestimmtes Puffersystem charakterisiert. Für den in dieser Arbeit verwendeten LightCycler ist das Protokoll für die Erstellung einer Kalibrierungsdatei zur Farbkompensation, welche die Parametermatrix enthält, im Benutzerhandbuch beschrieben. Die Fluoreszenzintensitäten der drei Farbstoffe Fluoreszein, RED-640 und RED-705 werden dazu im entsprechenden Puffer in getrennten Ansätzen bei unterschiedlichen Temperaturen gemessen. Die LightCycler-Software berechnet daraus die Parametermatrix und speichert sie in einer Datei („*color compensation calibration*“ mit der Endung „ccc“), die zur Farbkompensation in andere Messwertdateien importiert werden kann. Bei dieser Parametermatrix handelt es sich um die Parameter von insgesamt $4 \cdot 3 = 12$ Polynomen dritten Grades, welche die Signalverläufe der drei Fluorophore plus den einer Pufferkontrolle (ohne Fluorophor) in den drei Kanälen des LightCyclers in Anhängigkeit von der Temperatur beschreiben. Diese Interpretation der Matrix ergibt sich aus der vergleichenden Analyse der Daten aus der Messwertdatei zur Farbkompensation mit den Parametern der daraus erzeugten Kalibrierungsdatei.

Die Farbkompensation anhand dieser importierten Parametermatrix muss auch von der in dieser Arbeit entwickelten Software durchgeführt werden. Da es keine veröffentlichten Informationen zum Verfahren der Farbkompensation gibt, wurde in dieser Arbeit ein eigener Algorithmus zur Farbkompensation entwickelt. Er nutzt die Daten der von der LightCycler-Software erstellten Parametermatrix, kommt aber zu etwas anderen Ergebnissen als diese: Die Anteile des Übersprechens sind vollständig eliminiert, jedoch sind die residualen Messwerte nach der Farbkompensation geringer als die von der LightCycler-Software berechneten.

3.2.3.1 Prinzip

Das in einem Kanal K gemessene Fluoreszenzsignal S einer Mischung aus n unterschiedlichen Fluorophoren ist die Summe der Signalanteile S' aller n Fluorophore in diesem Kanal:

$$S_K = \sum_{i=1}^{n} S'_{K,i} \qquad \text{Gl. 3.16}$$

Die Indexvariable i bezeichnet die Nummer des Fluorophors. Die Signalanteile S' können als Teil a desjenigen Fluoreszenzwerts F beschrieben werden, den der Fluorophor in dem Kanal erzeugt, der sein Emissionsmaximum einschließt:

$$S'_{K,i} = a_{K,i} \cdot F_i \qquad \text{Gl. 3.17}$$

Durch diese Definition folgt für den Fall der Betrachtung des Kanals, der das Emissionsmaximum des betrachteten Fluorophors einschließt, dass der Signalanteil genau dem Fluoreszenzwert entspricht, also $a_{K,i} = 1$ (für $K = i$).

Eingesetzt in **Gl. 3.16** ergibt sich für die n Kanäle ($K = 1..n$) das lineare Gleichungssystem

$$S_K = \sum_{i=1}^{n} a_{K,i} \cdot F_i \qquad \text{Gl. 3.18}$$

bzw. in Matrixschreibweise

$$\boldsymbol{s} = \boldsymbol{A} \cdot \boldsymbol{f} \qquad \text{Gl. 3.19}$$

Der Vektor $\boldsymbol{s}$ enthält die gemessenen Fluoreszenzwerte aller Kanäle, $\boldsymbol{f}$ entsprechend die (kompensierten) Fluoreszenzwerte der Fluorophore im jeweiligen Kanal. Die Matrix $\boldsymbol{A}$ ($\boldsymbol{A} = a_{K,i}$ mit $K = 1..n$ und $i = 1..n$) lässt sich anhand der mit der Kalibrierung ermittelten Parametermatrix bestimmen und muss für jeden Temperaturwert der Messpunkte durchgeführt werden, da $\boldsymbol{A}$ temperaturabhängig ist. Für die Daten der Amplifikationskurven muss $\boldsymbol{A}$ nur einmal berechnet werden, da immer bei derselben Temperatur gemessen wird. Temperaturschwankungen von weniger als einem Kelvin bewirken nur vernachlässigbar kleine Unterschiede in $\boldsymbol{A}$. Daher braucht auch zur Auswertung von Schmelzkurven mit dem LightCycler $\boldsymbol{A}$ nur einmal je Messzyklus (der einmaligen Messung aller Proben) erfolgen: Die sequenzielle Messung von maximal 32 Proben erfordert im LightCycler etwa 5 s; bei einem Temperaturgradienten von 0.2 K/s beträgt der Temperaturunterschied der Messpunkte der ersten und letzten gemessenen Probe weniger als ein Kelvin.

3.2.3.2 Farbkompensation von Messwerten

Zur Berechnung der Kompensation wird zunächst aus den Werten der Parametermatrix eine Matrix von Polynomen dritten Grades erstellt. Die Matrixspalten entsprechen den Fluorophoren, die Matrixreihen den Kanälen. Für die im LightCycler verwendeten drei Fluorophore und drei Kanäle ergibt sich eine 3x3-Matrix. Das Element [1,1] bezieht sich auf den Fluorophor Fluoreszein, gemessen in Kanal 1, das Element [1,2] bezieht sich auf Fluoreszein, gemessen in Kanal 2 usw. Die Polynome beschreiben die in der Kalibrierungsmes-

sung ermittelten Signale der Fluorophore in den Kanälen in Abhängigkeit von der Temperatur. Anschließend wird anhand der Polynome eine Matrix von Datenreihen erstellt, welche die relativen, temperaturabhängigen Signalanteile der Fluorophore in den Kanälen beschreiben. Jede Datenreihe enthält 10 Datenpunkte für 10 unterschiedliche, gleichverteilte Temperaturwerte zwischen 40 °C und 100 °C. Die Werte dieser Datenreihen werden schließlich wieder durch Polynome interpoliert. Die nun vorliegende Matrix von Polynomen beschreibt die Signalanteile jedes in der Kalibrierungsmessung verwendeten Fluorophors in jedem Kanal für beliebige Temperaturen zwischen 40 °C und 100 °C.

Anhand dieser Polynommatrix wird jetzt für die Messtemperatur eine Matrix $\boldsymbol{A}$ berechnet, welche die Signalanteile der Fluorophore in den Kanälen für die gegebene Messtemperatur enthält. Die Bestimmung der kompensierten Messwerte erfolgt dann durch die Lösung des linearen Gleichungssystems **Gl. 3.19**.

Die Lösung dieser Gleichungssysteme muss im Prinzip für jeden Messpunkt durchgeführt werden und kann optimiert werden, indem $\boldsymbol{A}$ zunächst in die LU-Form (siehe unten) gebracht wird und das System anschließend nach dem CROUT-Algorithmus gelöst wird. Die Verfahren zur Dekomposition der Matrix und Lösung des Systems sind in Press *et al*. beschrieben.[218]

Die Dekomposition einer Matrix in die LU-Form leitet sich von der Darstellung einer Matrix als Produkt zweier triangulärer Matrizen ab: $A = L \cdot U$, *w*obei $\boldsymbol{L}$ die untere (*lower*) trianguläre und $\boldsymbol{U}$ die obere (*upper*) trianguläre Matrix ist. Im Falle einer 3x3-Matrix sieht diese Gleichung aus wie folgt:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \begin{bmatrix} \lambda_{1,1} & 0 & 0 \\ \lambda_{2,1} & \lambda_{2,2} & 0 \\ \lambda_{3,1} & \lambda_{3,2} & \lambda_{3,3} \end{bmatrix} \cdot \begin{bmatrix} \upsilon_{1,1} & \upsilon_{1,2} & \upsilon_{1,3} \\ 0 & \upsilon_{2,2} & \upsilon_{2,3} \\ 0 & 0 & \upsilon_{3,3} \end{bmatrix} \qquad \text{Gl. 3.20}$$

Mit dieser Form lässt sich ein lineares Gleichungssystem durch

$$\boldsymbol{A} \cdot \boldsymbol{f} = (\boldsymbol{L} \cdot \boldsymbol{U}) \cdot \boldsymbol{x} = \boldsymbol{L} \cdot (\boldsymbol{U} \cdot \boldsymbol{x}) = \boldsymbol{s} \qquad \text{Gl. 3.21}$$

lösen, indem zunächst der Vektor $\boldsymbol{y}$ bestimmt wird, so dass $\boldsymbol{L} \cdot \boldsymbol{y} = \boldsymbol{s}$ und dann $\boldsymbol{U} \cdot \boldsymbol{f} = \boldsymbol{y}$ gelöst wird.

Der Vorteil dieses Verfahrens liegt darin, dass diese beiden Systeme einen triangulären Satz linearer Gleichungen darstellen, die sich durch einfache Substitutionsverfahren sehr leicht lösen lassen. Während die aufwändige Berechnung der LU-Form nur einmal je Temperaturwert erfolgen muss (und damit z. B. nur ein einziges Mal für ein Experiment zur Quantifizierung), kann diese Form dann verwendet werden, um das System in **Gl. 3.19** sehr schnell für jeden Messpunkt jeder Probe zu lösen.

3.2.3.3 Vergleichende Auswertung mit dem neuen Algorithmus

Der Effekt der Farbkompensation ist im Folgenden an den Demonstrationsdateien zur parallelen Messung mit RED-640 und RED-705 gezeigt, die mit dem LightCycler ausgeliefert

werden. Die genauen Versuchsprotokolle zur Aufzeichnung der Daten sind dem LightCycler-Benutzerhandbuch zu entnehmen.

Das erste Beispiel zeigt Amplifikationskurven zweier unterschiedlicher Zielsequenzen im selben Reaktionsansatz (**Abb. 3-24**). Die erste Zielsequenz aus dem humanen *TNF*-Gen wurde mittels einer mit RED-640 markierten Hybridisierungssonde detektiert, deren Hauptsignalanteil in Kanal 2 aufgezeichnet wird. Bei der zweiten, nicht näher beschriebenen Zielsequenz handelt es sich um einen Kompetitor, der mit einer mit RED-705 markierten Hybridisierungssonde detektiert wurde, deren Hauptsignalanteil in Kanal 3 aufgezeichnet wird. Während die Proben unterschiedliche Mengen der *TNF*-Zielsequenz enthielten, war die Menge der Kompetitorsequenz in allen Proben gleich. Beide Zielsequenzen werden mit demselben Primerpaar amplifiziert.

Betrachtet man die Amplifikationskurven, die sich aus den in Kanal 2 aufgezeichneten Messwerten ergeben, erkennt man das typische Muster einer seriellen Verdünnungsserie der Zielsequenz mit äquidistanten exponentiellen Abschnitten der Kurven (**Abb. 3-24 A**).

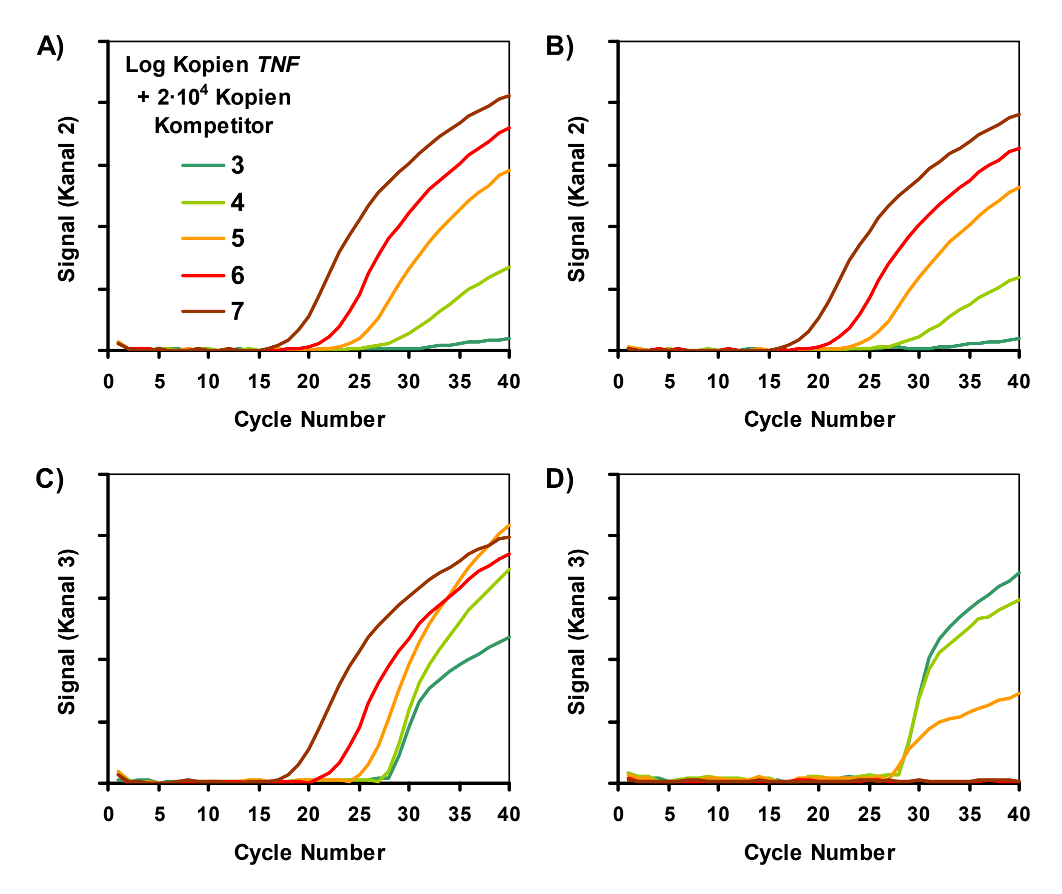


Abb. 3-24 Amplifikationskurven aus der Datei „DemoQuant-DualColor". **A)** Daten aus Kanal 2 (zur Detektion der *TNF*-Sequenz mit RED-640) ohne Farbkompensation. **B)** Daten aus Kanal 2 mit Farbkompensation. **C)** Daten aus Kanal 3 (zur Detektion der Kompetitorsequenz mit RED-705) ohne Farbkompensation. **D)** Daten aus Kanal 3 mit Farbkompensation.

Das Muster bleibt durch die Farbkompensation unverändert (**Abb. 3-24 B**).

Hierbei werden die Signalanteile des Fluoreszeins der Donorsonden sowie die des Farbstoffs RED-705 der Akzeptorsonden zur Detektion der Kompetitorsequenz vom gemessenen Signalwert in Kanal 2 subtrahiert. Wie aus **Abb. 2-3** auf Seite 39 ersichtlich, sind die in Kanal 2 detektierten Anteile der Emissionsspektren sowohl von Fluoreszein als auch von RED-705 sehr gering und haben einen entsprechend vernachlässigbaren Effekt bei der Farbkompensation in Kanal 2. Aufgrund der deutlich höheren Quantenausbeute des Fluoreszeins im Vergleich zu RED-640 ist der Signalanteil des Fluoreszeins in Kanal 2 jedoch messbar, ändert sich aber im Verlauf der PCR nicht signifikant, da die Sonden in einem großen Überschuss im Reaktionsansatz sind. Daher ist die durch die Zunahme des spezifischen PCR-Produkts bedingte Abnahme der Konzentration an freien Donorsonde während der Messung vernachlässigbar.

Deutlich anders verhält es sich mit den Amplifikationskurven, die sich aus den in Kanal 3 aufgezeichneten Messwerten ergeben. Ohne Farbkompensation erkennt man ein sehr ähnliches Muster wie für die Daten aus Kanal 2, obwohl die Amplifikationskurven wegen der in allen Reaktionsansätzen gleichen Menge an Kompetitor zumindest in ihren exponentiellen Phasen deckungsgleich sein sollten (**Abb. 3-24 C**). Das beobachtete Ergebnis lässt sich durch die signifikanten Signalanteile von RED-640 erklären, die in Kanal 3 gemessen werden: zum einen ist der in Kanal 3 detektierte Anteil des Emissionsspektrums von RED-640 vergleichsweise groß (**Abb. 2-3**), zum anderen ist die Quantenausbeute dieses Farbstoffs deutlich größer als die von RED-705. Nach der Farbkompensation zeigen nur noch die Amplifikationskurven der Proben mit weniger als 10^6 *TNF*-Kopien einen exponentiellen Signalanstieg, diesen aber, wie erwartet, alle beim selben Zykluswert (**Abb. 3-24 D**). In den Proben mit mehr als 10^5 *TNF*-Kopien, also mindestens der 5-fachen Menge des Kompetitors, wurde die Kompetitorsequenz nicht bis auf detektierbare Mengen amplifiziert, wahrscheinlich weil die im Reaktionsansatz vorhandenen Primer zur Amplifikation der *TNF*-Sequenz aufgebraucht wurden.

Das zweite Beispiel zeigt die Peakkurven einer Schmelzkurvenanalyse zur Detektion bzw. Typisierung zweier unterschiedlicher Mutationen auf dem gleichen Amplikon im selben Reaktionsansatz (**Abb. 3-25**). Das Versuchsprotokoll ist bei Bernard *et al*. Beschrieben.[219] Untersucht wurden die Kodone 112 und 158 des Gens für Apolipoprotein E (*apo* E) in einer Probe mit zwei Paaren von Hybridisierungssonden im Reaktionsansatz. Die Sequenz des Kodons 112 wurde mit einer RED-640-markierten Akzeptorsonde analysiert, die des Kodons 158 mit einer RED-705-markierten Akzeptorsonde. Ein beide Kodone umfassender Sequenzbereich wurde zunächst amplifiziert, anschließend wurde eine Schmelzkurve aufgezeichnet und anhand der ermittelten Schmelztemperaturen der Akzeptorsonden bestimmt, ob das Kodon die Wildtypsequenz aufweist („wt") oder eine Mutation trägt („mut"). Im Falle einer Mutation hybridisieren die zu den Wildtypsequenzen komplementären Akzeptorsonden nur unter Ausbildung einer Basenfehlpaarung auf dem Amplikon, welche die Schmelztemperatur des Hybrids herabsetzt.

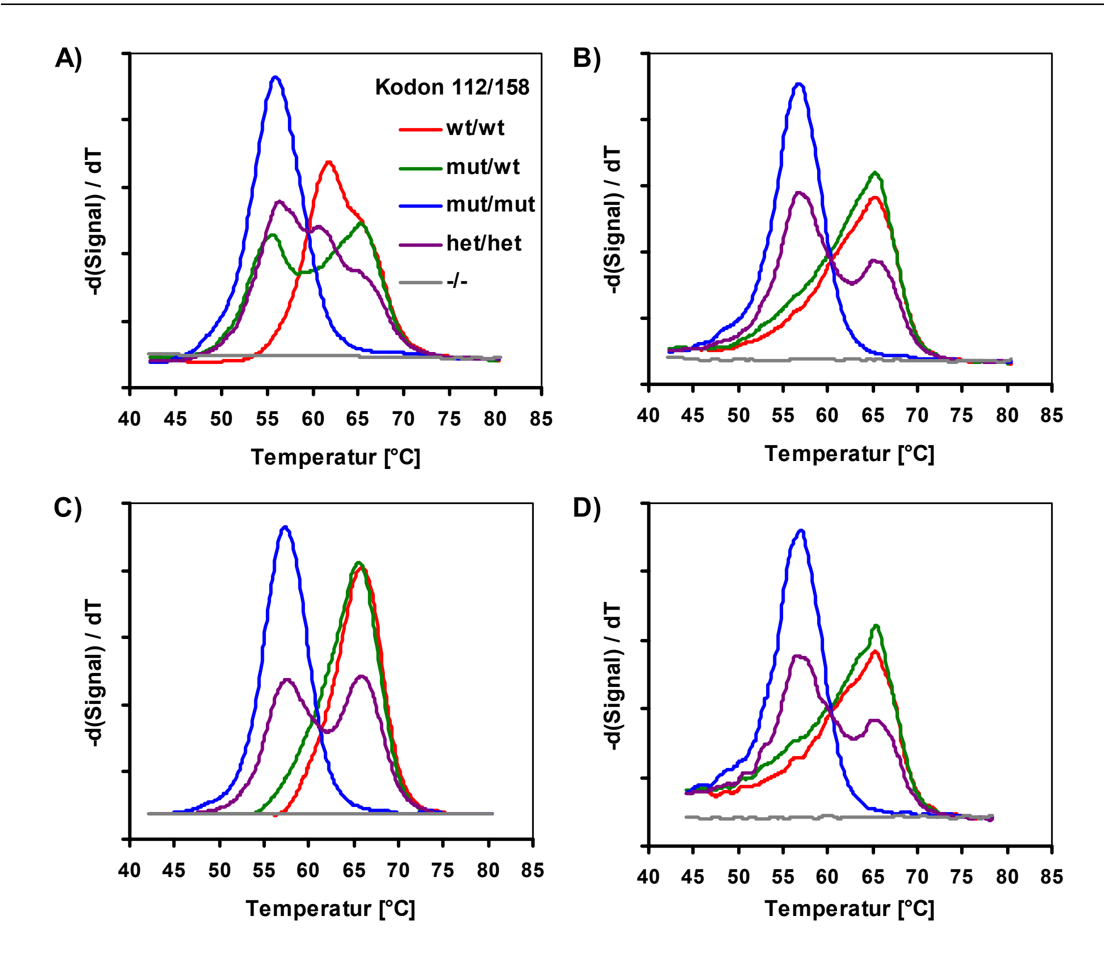


Abb. 3-25 Peakkurven der Schmelzkurvenanalyse der Daten aus „Demo-Mutation-DualColor". Verwendet wurden die Daten des Signalverhältnisses von Kanal 3 und Kanal 1: die Peakkurven sollen Aufschluss geben über den Genotyp in Kodon 158 des Gens *apo*E. Zur simultanen Bestimmung des Genotyps von Kodon 112 wurde eine mit RED-640 markierte Sonde verwendet, deren Schmelzverhalten sich aus dem hier nicht gezeigten Signalverhältnis von Kanal 2 und Kanal 1 ableiten lässt. **A)** Daten ohne Farbkompensation. **B)** Daten nach Farbkompensation. **C)** Daten nach Farbkompensation und Quenchkorrektur. **D)** Daten nach Farbkompensation, ausgewertet mit der LightCycler-Software LCDA, Version 3.5. wt/wt: homotygote Wildtypsequenz in beiden Kodonen, mut/wt: homozygote Mutation in Kodon 112, homozygote Wildtypsequenz in Kodon 158, mut/mut: homozygote Mutation in beiden Kodonen, het/het: heterozygot in beiden Kodonen, -/-: Negativkontrolle ohne Templat-DNA.

Die Peakkurve aus den in Kanal 3 (*RED-705, Kodon 158*) aufgezeichneten Daten der Probe $\frac{mut}{mut}/\frac{wt}{wt}$ (homozygote Mutation in Kodon 112, homozygote Wildtypsequenz in Kodon 158) sollte genau einen Peak haben mit dem Maximum bei ca. 66 °C, zeigt jedoch zwei Peaks, was eigentlich einen heterozygoten Genotyp für Kodon 158 anzeigen würde. Bei der Probe $\frac{mut}{wt}/\frac{mut}{wt}$ (heteozygot an beiden Positionen) sollten hier hingegen zwei Peaks erkennbar sein mit den Maxima bei ca. 57 °C und 66 °C. Stattdessen zeigt die Peakkurve

drei überlagerte Peaks. Lediglich die Kurven für die Proben $\frac{wt}{wt}/\frac{wt}{wt}$ und $\frac{mut}{mut}/\frac{mut}{mut}$ können richtig und zweifelsfrei interpretiert werden (**Abb. 3-25 A**). Nach Farbkompensation lassen sich alle Kurven korrekt interpretieren (**Abb. 3-25 B**). Wie in **Abb. 3-25 C** zu sehen, wird das Ergebnis noch deutlicher, insbesondere für die Probe $\frac{mut}{wt}/\frac{mut}{wt}$, wenn die Daten zusätzlich um den temperaturabhängigen Quench korrigiert werden (siehe **3.2.2.2**). **Abb. 3-25 D** zeigt zum Vergleich die mit der LightCycler-Software LCDA (Version 3.5) ermittelten Peakkurven nach Farbkompensation. Das Ergebnis ist vergleichbar mit den in **B** gezeigten Kurven, lediglich die Glättung der Kurven ist hier etwas schlechter (die Kurven aus **A**, **B** und **C** sind wie in **0** beschrieben durch *Splines* geglättet).

3.2.4 Software zur automatisierten Auswertung

Die in **3.2.1**, **0** und **3.2.3** vorgestellten Algorithmen wurden in ein Computerprogramm implementiert, welches die komfortable und automatisierte Auswertung von Real-time-PCR-Messungen ermöglicht. Zusätzlich werden die Ergebnisse von Quantifizierungen hinsichtlich ihrer Genauigkeit und Verlässlichkeit charakterisiert. Das Programm erlaubt zudem die quantitative Auswertung von Experimentdateien nach den herkömmlichen Verfahren (ohne Korrektur amplifikationsunabhängiger Trends aus den Signalkurven und Bestimmung der C_T-Werte nach der *Fit-points*-Methode oder der *2nd-derivative-maximum*-Methode, wie sie im LightCycler Benutzerhandbuch beschrieben ist). Der Programmname ist *S*oFAR, was sich ableitet von *So*ftware *F*or the *A*nalysis of *R*eal-time PCR data. Das Programm ist vollständig objektorientiert programmiert. Als Entwicklungsumgebung diente Borland Delphi in der Version 6.0.

Im Rahmen dieser Arbeit kann und soll nicht auf die Details der Programmierung eingegangen werden. Der dokumentierte Quelltext sowie das lauffähige Programm befinden sich auf der beiliegenden CD-ROM im Verzeichnis „SoFAR". Hier wird im Folgenden nur ein Überblick über die Konzeption, und die Verbindungsstrukturen der an den Auswertungen beteiligten Klassen gegeben. Auf die Komponenten zur Anzeige von Datenreihen in x/y-Diagrammen (sehr umfangreiche und leistungsfähige Eigenentwicklungen) beispielsweise, wird hier nicht eingegangen. Zwar ist die grafische Präsentation der Daten ein wesentlicher Teil der Software, allerdings haben diese Komponenten keinen Anteil an der Auswertung von Messwertdateien.

Bezeichner von Programmteilen wie Objekten, Variablen und Prozeduren, wie sie im Programmcode verwendet werden, sind durch die Schriftart `Courier` gekennzeichnet. Das Programm selbst ist mit dem Quellcode im Anhang hinreichend kommentiert.

Die zum Verständnis der folgenden Abschnitte wichtigen programmiertechnischen Begriffe sowie Bezeichnungen, die im Zusammenhang mit RT-PCR-Experimenten stehen, sind im Anhang in **Tab. 7-1** erklärt (**7.3**), Beschreibungen der wichtigen Klassen und Datenfelder sind ebenfalls im Anhang (**7.4**) aufgeführt.

3.2.4.1 Konzeption

Mit dem Softwareprogramm ist eine parallele Auswertung von Proben verschiedener Experimente möglich, sofern die PCR-Protokolle dieser Experimente genau übereinstimmen. Dies erlaubt z. B. den direkten Vergleich der Ergebnisse verschiedener Experimente, die nach demselben Protokoll durchgeführt wurden, oder die Quantifizierung von Proben aus mehreren Experimenten anhand einer zuvor aufgezeichneten Standardkurve. Der direkte Vergleich bzw. die gemeinsame quantitative Auswertung von Experimenten mit unterschiedlichen PCR-Protokollen ist in der Regel nicht sinnvoll und wird daher vom Programm nicht erlaubt. Es besteht jedoch die Möglichkeit, dass Proben gemeinsam ausgewertet werden, bei welchen unterschiedliche Zielsequenzen amplifiziert wurden und/oder unterschiedlich zusammengesetzte Reaktionsansätze verwendet wurden. Darüber liegen im PCR-Protokoll keine Informationen vor, so dass hier immer der Anwender sicherstellen muss, dass nur sinnvolle Kombinationen von Proben gemeinsam ausgewertet werden.

Allgemeine Informationen zu den geladenen Experimenten sowie die Ergebnisse der Analysen werden in tabellarischer Form dargestellt, die zur Auswertung verwendeten Datenreihen mit Messwerten werden in x/y-Diagrammen dargestellt. Die Anzeige und Auswertung bezieht sich auf eine Selektion von Proben der geladenen Experimente, die jederzeit vom Anwender geändert werden kann, um z. B. nur eine Untergruppe von Proben auszuwerten oder um Ausreißer aus der Auswertung herauszunehmen.

Wenngleich normalerweise nicht notwendig, so erlaubt das Programm dennoch die Änderung aller wichtigen Einstellungen für die Analysen wie z. B. der zu verwendenden Messwerte, unterschiedlicher Korrekturverfahren der Rohdaten, Schwellenwerte, unterschiedlicher Verfahren zur Bestimmung der C_T-Werte usw. Jede Änderung bewirkt die augenblickliche Neuberechnung aller Ergebnisse, so dass jederzeit jede Einstellung geändert werden kann. Der Anwender wird nicht, wie bei den derzeit kommerziell erhältlichen Programmen, schrittweise durch die Auswertung geführt, wo in jedem Schritt ein anderer Analyseparameter einzustellen und zu kontrollieren ist, sondern die Ergebnisse der Analysen liegen immer vor und sind immer gültig für die aktuell selektierten Proben und die aktuellen Einstellungen. Bei Experimenten mit eindeutig auswertbaren Daten braucht der Anwender das Experiment nur zu öffnen, um sofort und ohne weitere Angaben das Ergebnis aller Analysen zu erhalten. Bei zweifelhaften Ergebnissen kann der Anwender jede Einstellung direkt verändern und erhält sofort die neuen Ergebnisse.

Der Quellcode des Projekts gliedert sich in verschiedene Abschnitte, die `Units`, welche jeweils logisch zusammenhängende Teile des Quellcodes enthalten. Diese Gliederung des Quelltextes erhöht die Übersichtlichkeit des Projekts und erlaubt eine schnelle und einfache Navigation. Die Namen der Quellcodedateien selbsterstellter `Units` beginnen mit einem „U“, um sie als `Unit`-Quellcodedatei zu kennzeichnen. Eine Liste der für dieses Projekt erstellten `Units` ist in **Tab. 3-1** aufgeführt. Die zur Auswertung wichtigen Strukturen aus diesen `Units` sind in den folgenden Abschnitten behandelt.

Die Units stehen in einem hierarchischen Abhängigkeitsverhältnis zueinander, da sie Teile aus anderen Units benötigen (zirkuläre Abhängigkeiten sind in Delphi nicht erlaubt). Beispielweise benötigt die in der Unit UAnalysis definierte Klasse TAnalysisData einen Zugriff auf die Messdaten, die von der Klasse TSample verwaltet werden, die ihrerseits in der Unit TRun definiert ist. Dazu muss die Unit UAnalysis die Unit URun einbinden. Änderungen innerhalb der Unit URun können somit Auswirkungen auf die Funktionsfähigkeit der in der Unit UAnalysis definierten Klassen haben. Eine Übersicht der Abhängigkeiten ist in **Abb. 3-26** gezeigt, aus welcher auch funktionale Gruppen von Units hervorgehen, die durch einheitliche Farbgebungen gekennzeichnet sind.

Tab. 3-1 **Dateinamen und Inhalte der Units, die im Projekt zur Datenverwaltung und Analyse verwendet werden.** Die meisten Units beinhalten die Deklaration von einer oder mehreren Klassen. Die Bezeichner der wichtigsten Klassen sind jeweils im Inhaltsfeld angegeben.

Dateiname	Inhalt	Klassen
Units für fundamentale Aufgaben		
UArrays	Klassen für dynamische Felder zur Verwaltung von Datenreihen und Matrizen variabler Größe.	TFltList u. a.
UFlatComponents	Standardkomponenten wie Schalt- und Auswahlfelder in flachem 3D-Design.	TFlatEdit TFlatCombo u. a.
UObjList	Klassen zur Verwaltung von Objektlisten.	TObjList TListObj
UParser	Klasse zum Einlesen und zur Interpretation von Textdateien.	TParser
UStream	Klassen zum Speichern und Laden veröffentlichter Eigenschaftswerte von Objekten.	TPersistentWriter TPersistentReader TStreamObject
UUserDialog	Abstrakte Dialogkomponente zur Darstellung selbstdefinierter Dialoge.	TUserDialog
Units für mathematische Funktionen		
UFit	Funktionen zur linearen Regression sowie eine Klasse zur Verwaltung parametrischer Funktionen, deren Parameter an gegebene Datenreihen angepasst werden können.	TFitFunc
UFunctions	Klassen zur Verwaltung verschiedener Funktionen.	TBgrFunction TExponential u. a.

Dateiname	Inhalt	Klassen
UMath	Grundlegende mathematische und statistische Funktionen zur Berechnung von Logarithmen, Potenzen, Mittelwerten, Standardabweichungen und Ableitungen sowie Klassen zur Verwaltung von mathematischen Funktionen ohne und mit Parametern.	TFunc TParamFunc
UPeaks	Klasse zur Verwaltung der Peakfunktion (Summe von Gaußkurven).	TPeaks
USpline	Klasse zur Verwaltung von *Spline*-Funktionen.	TSpline
Units zur Datenanalyse		
UAnalysis	Abstrakte Komponente zur Analyse von Experimenten. Die Klassen werden in den Units UQuantAnalysis und UMeltAnalysis abgeleitet.	TAnalysis TAnalysisData TAnalysisGroup
UMeltAnalysis	Komponente zur Schmelzkurvenanalyse von Experimenten sowie Klassen zur Verwaltung der Daten und Ergebnisse der analysierten Proben und Probengruppen.	TMeltAnalysis TMeltAnalysisData TMeltAnalysisGroup
UQuantAnalysis	Komponente zur quantitativen Analyse von Experimenten sowie Klassen zur Verwaltung der Daten und Ergebnisse der analysierten Proben und Probengruppen.	TQuantAnalysis TQuantAnalysisData TQuantAnalysisGroup
Units zur grafischen Anzeige von Datenreihen		
UChart	Visuelle Komponente zur Darstellung von x/y-Diagrammen.	TChart TChartData u. a.
UChartComponents	Komponenten zur Einstellung verschiedener Eigenschaften von Diagrammelementen der in UChart deklarierten Komponente TChart.	TChartCombo u. a.
UChartDialog	Dialog zur Einstellung der Eigenschaften von Diagrammelementen der in UChart deklarierten Komponente TChart.	TChartDialog
UDataStyleDialog	Dialog zur Einstellung der Eigenschaften zur Anzeige von Datenreihen in Diagrammen der in UChart deklarierten Komponente TChart.	TDataStyleDialog

Dateiname	Inhalt	Klassen
Units zur Datenverwaltung		
ULCTreeView	Visuelle Komponente zur Verwaltung von LightCycler-Experimenten und zur Anzeige und Selektion von Proben.	TLCTreeview
URun	Klassen zur Verwaltung der Daten von LightCycler-Experimenten.	TRun TSample u. a.

Die drei wichtigsten Gruppen umfassen (i) Units für mathematische Berechnungen wie der Berechnung von Funktionswerten und statistischen Kennwerten sowie der Anpassung von Parametern, (ii) Units zur grafischen Darstellung von Datenreihen in x/y-Diagrammen sowie Dialogen zur Einstellung von Anzeigeattributen und (iii) Units zur A-nalyse der Datenreihen für quantitative Auswertungen oder Schmelzkurvenanalysen. Die Units URun und ULCTreeview lassen sich keiner dieser Themengruppen zuordnen, wobei URun eine zentrale Stellung einnimmt. Sie bindet UFunctions ein, da die Klasse TPolynomial zur Farbkompensation bei Multiplexexperimenten benötigt wird (siehe **3.2.3**). ULCTreeview wiederum verbindet die Datenverwaltung über Instanzen der in URun definierten Klasse TRun mit den Anzeigeattributen der Datenreihen in den x/y-Diagrammen, weshalb sie neben URun auch Units aus den Themengruppen zur grafischen Darstellung einbindet.

Auf der obersten Ebene finden sich die Units ULCTreeview, UQuantAnalysis und UMeltAnalysis, welche die zentralen Komponenten mit den zugehörigen Klassen zur Datenverarbeitung bereitstellen, die im Hauptprogramm genutzt werden.

3.2.4.2 Klassen zur Datenverwaltung und Auswertung

Eine Kurzbeschreibung der Klassen zur Verwaltung und Auswertung von LightCycler-Daten befindet sich in **Tab. 7-2** (**7.4**), eine Übersicht über den Stammbaum dieser Klassen ist in **Abb. 7-1** (**7.4**) gezeigt.

Die Verwaltung der vom LightCycler-System aufgezeichneten Daten geschieht im Wesentlichen über die Klasse TRun, welche die Daten aus den Dateien lesen kann und anderen Objekten für Auswertungen und Anzeigen zur Verfügung stellt. Instanzen der Klasse TRun werden von der Klasse TLCTreeView verwaltet, welche sicherstellt, dass die Protokolle gleichzeitig geöffneter Dateien übereinstimmen sowie eine Schnittstelle zur Selektion von Proben durch den Anwender bereitstellt.

Die Berechnungen zur Quantifizierung und Schmelzkurvenanalyse werden von den von TAnalysis abgeleiteten Klassen durchgeführt. Die dazu benötigten probenbezogenen Daten wie z. B. Angaben über die exponentiellen Phasen von Amplifikationskurven oder über die Peaks von Schmelzkurvendaten werden von den von TAnalysisData abgeleiteten Klassen verwaltet. Die von TAnalysisGroup abgeleiteten Klassen sind zuständig für die statistischen Auswertungen von Probengruppen.

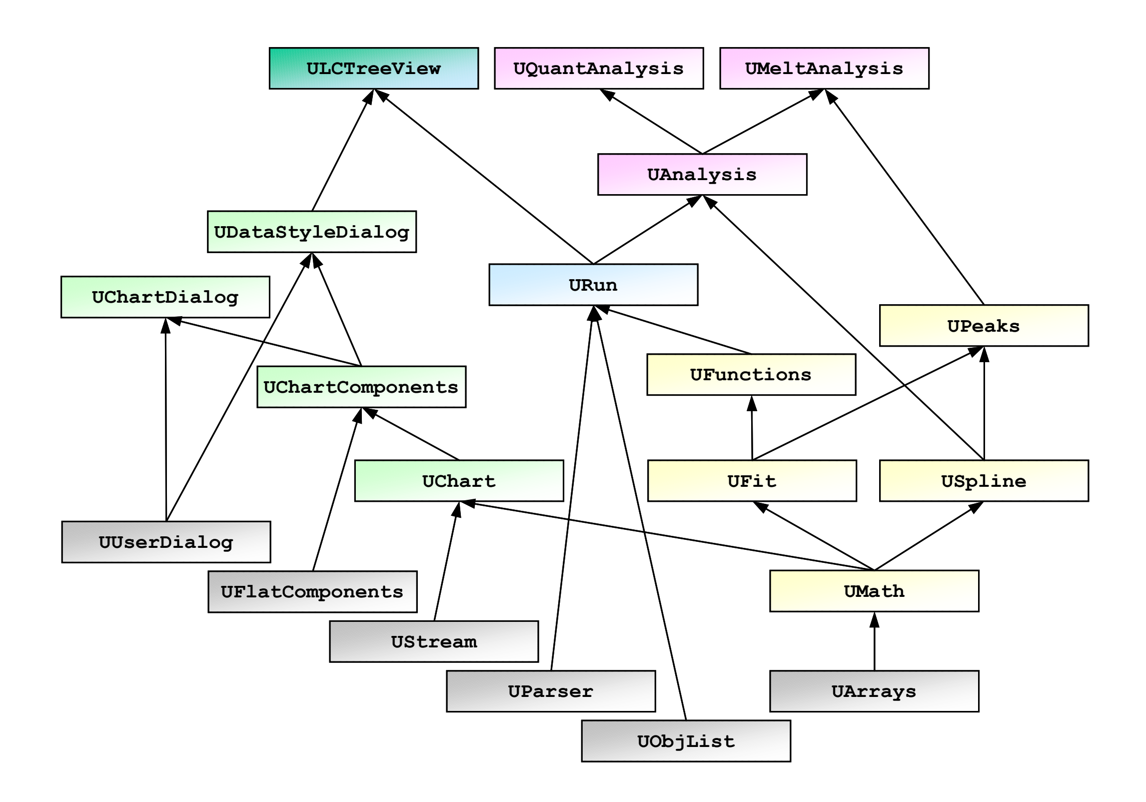


Abb. 3-26 Abhängigkeitsdiagramm der selbsterstellten Projekt-`Units`. Die `Units` der höheren Ebenen binden die durch Pfeile verbundenen `Units` niedrigerer Ebenen ein. Die `Units` binden immer alle `Units` ein, von denen die eingebundenen `Units` abhängig sind. `URun` z. B. bindet auch `UArrays` ein, was aus der Reihe `UArrays–UMath–UFit–UFunctions–URun` hervorgeht. Die Farben kennzeichnen zusammengehörige Themengruppen (grau: fundamentale Aufgaben, grün: grafische Anzeige, gelb: mathematische Funktionen, rot: Analyse, blau: Datenverwaltung).

3.2.4.3 Klassenstruktur zur Verwaltung der LightCycler-Daten

Die vom LightCycler-System aufgezeichneten Daten beziehen sich auf jeweils ein Experiment. Im *„About-File“* (.ABT) befinden sich Informationen zum Experiment, zu Geräteeinstellungen, zum PCR-Protokoll sowie zu den analysierten Proben. Die aufgezeichneten Messwerte für Zeit, Temperatur und Fluoreszenz befinden sich im *„Fluorescence-File“* (.FLO). Für die Auswertungen werden die Daten aus diesen Dateien eingelesen und in strukturierter Form über ein Objekt der Klasse `TRun` bereitgestellt, welches seinerseits Schnittstellen zu anderen Objekten bereitstellt, welche spezielle Daten verwalten.

Eine Übersicht über die beteiligten Klassen, deren Daten sowie die Verbindungen zwischen den Klassen ist in **Abb. 3-27** dargestellt. Die Bezeichnungen der Datenfelder mit Angaben über die Datentypen und deren Bedeutung sind in der Tabelle **Tab. 7-3** (**7.4**) aufgeführt.

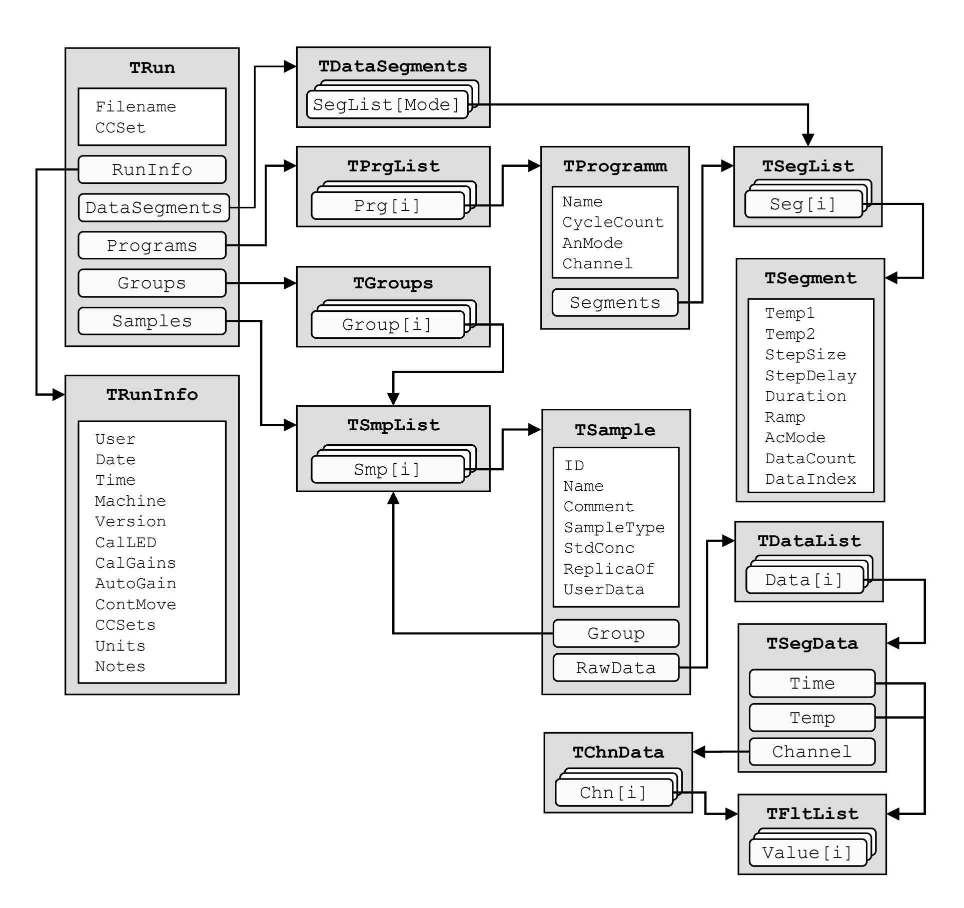


Abb. 3-27 Klassenstruktur zur Verwaltung von LightCycler-Daten. Die Klassenstrukturen sind durch graue Rechtecke symbolisiert. Sie enthalten den Klassenbezeichner (fett), Verweise auf andere Objekte (Pfeile) sowie ggf. Feldbezeichner bzw. Eigenschaften (in den weißen Kästen) für die gespeicherten Daten. Die Bedeutungen der Felder sind in Tab. 7-3 beschrieben. Methoden und Ereignisse der aufgeführten Klassen sind der Übersichtlichkeit halber nicht angegeben.

Die Klasse `TRun` speichert selbst lediglich den Dateinamen der geladenen LightCycler-Datei (`Filename`) sowie die Parameter der verwendeten Farbkompensation (`CCSet`), da die Farbkompensation der Fluoreszenzdaten der Proben von `TRun` durchgeführt wird. Alle anderen probenübergreifenden Daten werden in der Klasse `TRunInfo` gespeichert. Das PCR-Protokoll ist über den Eintrag `Programs` verfügbar, der auf eine Liste mit PCR-Programmen zeigt, die ihrerseits wieder Verweise auf Listen mit den Segmentdaten enthalten. Die probenbezogenen Daten sind über `Samples` erreichbar. Probengruppen identischer Reaktionsansätze sind durch `Groups` ansprechbar. Die Eigenschaft `RawData` Klasse `TSample` verzweigt zu den Messwertdaten der Probe. Die dort referenzierte Klasse `TDataList` enthält eine Liste mit Verweisen auf alle Messwertlisten der Probe, die je Segment in einer Klasse `TSegData` zusammengefasst sind. Der Eintrag im Feld

`DataIndex` der Klasse `TSegment` ist der Index desjenigen Eintrags in `TDataList`, welcher die Daten dieses Segments enthält.

Die Klasse `TDataSegments` enthält eine nach dem Analysemodus (Quantifizierung, Schmelzkurve) sortierte Aufstellung von Listen mit Segmenten, welche Daten enthalten. In der Liste `TRun.DataSegments[anQuantification]` befinden sich z. B. alle Segmente des PCR-Protokolls, die Daten zur Quantifizierung enthalten.

Die Datenreihen mit den Zeit-, Temperatur- und Fluoreszenzwerten werden von Objekten der Klasse `TFltList` verwaltet. Der benötigte Speicherplatz für die Fließkommawerte wird je nach Bedarf dynamisch alloziert. Zwar bietet die verwendete Programmiersprache dynamische Feldvariablen, die genau diesen Zweck erfüllen („*dynamic arrays*") , jedoch funktionieren diese in der aktuellen Version der Entwicklungsumgebung nicht fehlerfrei. Daher war die Neuentwicklung einer Klasse zur dynamischen Speicherverwaltung nötig.

Die Objekte der Klasse `TRun` werden von einem übergeordneten Objekt verwaltet, welches eine visuelle Schnittstelle zur Auswahl der auszuwertenden Proben zur Verfügung stellt. Bei diesem Objekt, `TLCTreeView`, handelt es sich um eine von `TTreeView` abgeleitete Komponente, welche in der Lage ist, die Proben der geladenen Experimente in einer baumartig verzweigten Struktur darzustellen. Die Einträge in dieser Struktur werden als Knoten bezeichnet. Jeder Knoten kann seinerseits wieder eine beliebige Zahl untergeordneter Knoten enthalten. Die Knoten der obersten Ebene repräsentieren die geladenen Experimente. Jeder Experiment-Knoten verzweigt in drei Unterknoten, welche die Proben gleichen Probentyps enthalten: Standards, Unbekannte und Kontrollen. Wenn Proben gruppiert wurden, werden diese auch in `TLCTreeView` unter einem Gruppen-Knoten zusammengefasst.

Die Komponente `TLCTreeView` verwaltet die Daten über eigene Datenstrukturen, welche Verweise auf die Lauf-, Gruppen- oder Probendaten aus `TRun` mit Verweisen auf Darstellungsparameter für die grafische Anzeige und Informationen zum Selektionsstatus der Probe zusammenfassen.

Die Anforderung, ein Experiment zu öffnen oder zu schließen, erfolgt im Programm über die Komponente `TLCTreeView`. Diese prüft zunächst, ob die gewünschte Aktion, z. B. das Hinzufügen eines Experiments zur Auswertung, möglich ist. Wenn das der Fall ist, wird die Aktion ausgeführt, ansonsten wird ein Fehlerwert zurückgegeben. Wurde eine Aktion ausgeführt, erzeugt die Komponente ein entsprechendes Ereignis, worauf andere Komponenten reagieren können, um nötige Änderungen vorzunehmen (z. B. die Aktualisierung von Ergebnissen und Anzeigen).

Die gemeinsame Auswertung von Proben ergibt nur dann sinnvolle Ergebnisse, wenn sowohl die Reaktionskomponenten (bis auf das Templat) als auch die PCR-Programme bei allen gemeinsam auszuwertenden Proben identisch sind. Da sich Angaben zu PCR-Programmen in den LightCycler-Dateien befinden, kann ein Experiment nur dann bereits geöffneten Experimenten hinzugefügt werden, wenn die PCR-Programme in allen Punkten übereinstimmen. Über die Reaktionskomponenten sind in den Dateien keine zwingenden

Angaben gemacht. Daher muss der Anwender sicherstellen, dass alle gleichzeitig selektierten Proben, egal ob aus einem oder mehreren Experimenten, identische Reaktionskomponenten beinhalten.

3.2.4.4 Klassenstruktur zur Auswertung selektierter Proben

Die Auswertung der selektierten Proben erfolgt durch die eigenständigen Komponenten `TQuantAnalysis` und `TMeltAnalysis`, die von der Komponente `TAnalysis` abgeleitet sind. `TAnalysis` verweist auf ein Segment (`TSegment`), dessen Daten analysiert werden sollen. Weiterhin verwaltet sie eine Liste mit probenbezogenen Daten sowie eine Liste mit gruppenbezogenen Daten. Die probenbezogenen Daten werden in der Klasse `TAnalysisData` verwaltet, die gruppenbezogenen Daten in der Klasse in `TAnalysisGroup`. `TAnalysis` führt die Methode `Update` ein, in welcher die virtuellen Methoden `CalcData` und `CalcResults` sowie die Methoden zur Ereignisgenerierung aufgerufen werden.

Eine Übersicht über die beteiligten Klassen, deren Daten sowie die Verbindungen zwischen den Klassen ist in **Abb. 3-28** dargestellt. Die Bezeichnungen der Datenfelder mit Angaben über die Datentypen und deren Bedeutung sind in **Tab. 7-4** (**7.4**) aufgeführt.

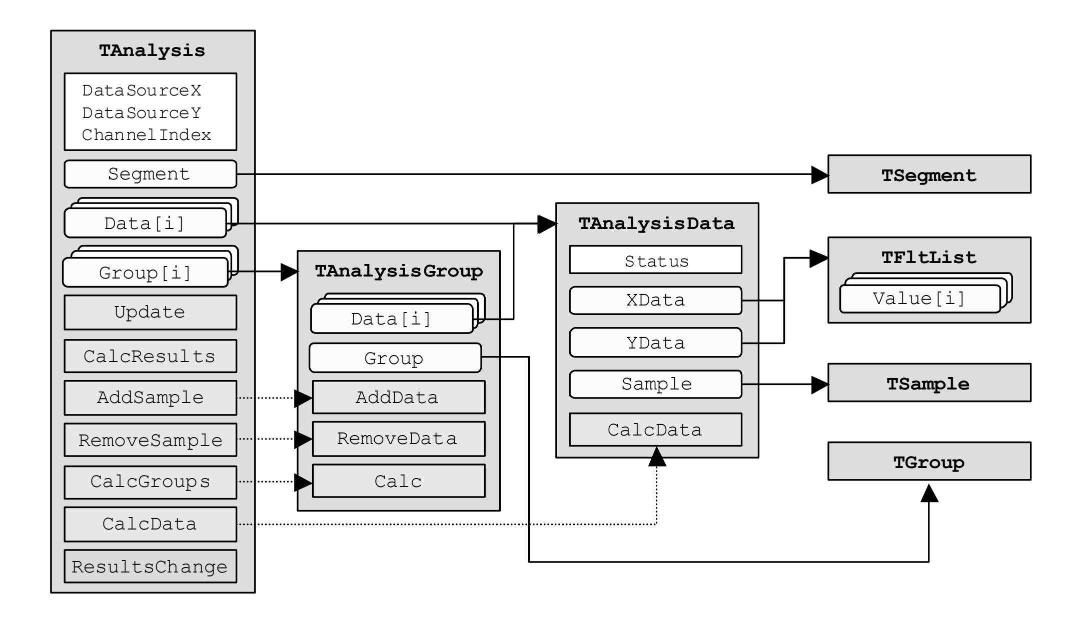


Abb. 3-28 Klassenstruktur zur Analyse der selektierten Proben. Die Klassen `TAnalysis`, `TAnalysisData` und `TAnalysisGroup` sind abstrakte Klassen. Im Programm werden die speziell für die Quantifizierung und Schmelzkurvenanalyse abgeleiteten Klassen verwendet, bei welchen die virtuellen Methoden der abstrakten Klassen überschrieben sind. Außer der Methode `Update` sind alle hier aufgeführten Methoden virtuell deklariert.

Die Klasse `TAnalysisData` enthält einen Verweis auf die Probe (`TSample`) sowie Listen mit Fließkommawerten (x- und y-Werte) der zu analysierenden Daten. Die Klasse hat einen virtuellen Konstruktor und eine ebenfalls virtuell deklarierte Methode `CalcData`, die in den abgeleiteten Klassen `TQuantAnalysis` und `TMeltAnalysis` zur Quantifizierung bzw. zur Schmelzkurvenanalyse überschrieben sind. In der Methode `TAnalysisData.CalcData` werden lediglich die x- und y-Werte aus den Messdaten zugewiesen. Die überschriebenen Methoden ergänzen dies um analysespezifische Berechnungen. Die Berechnungen im Einzelnen erfolgen gemäß der in **3.2.1** und **0** beschriebenen Algorithmen.

3.2.4.5 Klassen zur Verwaltung von Funktionen

Datenreihen zur quantitativen Analyse sowie zur Schmelzkurvenanalyse werden durch verschiedene Funktionen beschrieben (siehe **3.2.1** und **0**). Die Parameter dieser Funktionen werden zusammen mit den Methoden zur Berechnung der Funktions- und Ableitungswerte von speziellen Klassen verwaltet. Bei allen Funktions- und Parameterwerten handelt es sich um Fließkommazahlen doppelter Genauigkeit (Datentyp `Double`: 8 Byte, 15 signifikante Stellen). Die von `TObject` abgeleitete, abstrakte Klasse `TFunc` stellt die abstrakte Funktion `y(x)` zur Verfügung, die in abgeleiteten Klassen überschrieben werden muss, um anhand des angegebenen x-Wertes einen y-Wert zu berechnen. Die virtuelle Funktion `dy(x)` berechnet die erste Ableitung bei x numerisch nach der Formel $dy(x) = (y(x+h)-y(x))\cdot h^{-1}$., wobei h ausgehend von 10^{-8} solange um den Faktor 0.2 verringert wird, bis sich der berechnete Wert für $dy(x)$ nicht mehr ändert oder nicht mehr konvergiert, die Änderung also wieder zunimmt. Mit der ebenfalls virtuell deklarierten Funktion `ddy(x)` kann, wie bereits für `dy(x)` beschrieben, die zweite Ableitung der Funktion `y(x)` numerisch ermittelt werden, wobei hier $ddy(x) = (y(x+h+h)\text{-}2\cdot y(x+h)+y(y))\cdot h^{-2}$. Die von `TFunc` abgeleitete Klasse `TParamFunc` ermöglicht die Verwaltung von Parametern und stellt die dazu notwendigen Methoden zur Verfügung.

Die Klasse `TFitFunc`, ihrerseits abgeleitet von `TParamFunc`, implementiert die virtuelle Methode Fit, über welche die Parameter der Funktion so angepasst werden, dass eine gegebene Koordinatenreihe möglichst gut durch die Funktion beschrieben wird. Dazu wird die Summe der Fehlerquadrate zwischen Funktions- und Datenwerten (χ^2) minimiert (*least-squares fit*). Die numerische Lösung dieses Problems erfolgt über den LEWENBERG-MARQUARD-Algorithmus[220]. Dabei wird, ausgehend von Schätzwerten für die Parameter, anhand der partiellen Ableitungen von χ^2 nach den Parametern der Funktion ein Gradient ermittelt (die sogenannte HESSE-Matrix), der angibt, wie die Parameter verändert werden müssen, um zu einem kleinern χ^2 zu gelangen. Nach der Änderung der Parameter und Berechnung des neuen χ^2 wird diese Prozedur solange wiederholt, bis χ^2 konvergiert bzw. sich nicht mehr signifikant verkleinert. Die Schrittweite der Änderungen wird für jeden Schritt neu angepasst. Wenn sich χ^2 nach einem Schritt verschlechtert, wird die Schrittweite vergrößert, im anderen Fall wird sie verringert. Einen Anhaltspunkt für die Größenordnung der zu wählenden Schrittweite erhält man bei diesem Verfahren durch die Formulierung

des Gradienten von χ^2 (HESSE-Matrix). Darüber kann das jeweils zu lösende Gleichungssystem so formuliert werden, dass die Schrittweite als Faktor angegeben werden kann.

Die Anpassungen der Parameter nichtlinearer Funktionen durch die Minimierung von χ^2 ist auch nach dem bekannten Simplex-Verfahren möglich. Simplex bezeichnet einen Körper in einem n-dimensionalen Raum mit der minimal möglichen Anzahl an Eckpunkten ($n+1$), z. B. einen Tetraeder im dreidimensionalen Raum. Die Anzahl n der Dimensionen des Raums entspricht hier der Anzahl der Parameter der zu minimierenden Funktion. Zum Start wird ein Simplex durch n Punkte um einen Startwert herum definiert. Zu jedem Punkt wird dann χ^2 berechnet. Der Punkt mit dem größten χ^2 wird an der Seite gespiegelt, deren Eckpunkte die kleinsten χ^2-Werte aufweisen. Die Spiegelung erfolgt mit einem variablen Faktor, der so groß gewählt wird, dass der χ^2-Wert des gespiegelten Eckpunkts nicht größer wird. Im Gegensatz zum oben beschrieben LEWENBERG-MARQUARD-Algorithmus brauchen hier keine partiellen Ableitungen berechnet werden (die für bestimmte Parameterwerte nicht definiert sein könnten), jedoch müssen hier für eine vernünftige Wahl von Startwerten und Schrittweiten die Größenordnungen für sinnvolle Parameterwerte und die „Glattheit" des χ^2-Raumes bekannt sein. Der LEWENBERG-MARQUARD-Algorithmus lässt sich daher leichter zur Anpassung der Parameter sehr unterschiedlicher Funktionen an unterschiedlichste Datenreihen verwenden als die Simplex-Methode.

Die zur Glättung und Analyse von Koordinatenreihen verwendeten *Splines* werden durch die von `TFunc` abgeleitete Klasse `TSpline` verwaltet. Wie `TFitFunc` stellt auch `TSpline` eine Methode `Fit` zur Verfügung, welche die Bestimmung der Parameter zur Beschreibung einer Koordinatenreihe durchführt. Bei den Parametern handelt es sich hier um Listen von Parametersätzen mit je 4 Elementen, welche je ein kubisches Polynom definieren. Die Intervallgrenzen der Polynome sind in einer weiteren Parameterliste gespeichert. Zur Berechnung von Funktions- und Ableitungswerten muss zunächst das Intervall bestimmt werden, welches den angegebenen x-Wert enthält, damit der entsprechende Parametersatz identifiziert werden kann, mit dem der Funktionswert zu berechnen ist. Die virtuellen Methoden `dy` und `ddy` von `TFunc` zur Berechnung der ersten und zweiten Ableitung sind überschrieben, da diese sich hier direkt aus den Parametern der Polynome berechnen lassen.

3.2.4.6 Charakterisierung der Quantifizierungsergebnisse

Die Ergebnisse der Quantifizierung werden hinsichtlich ihrer Gültigkeit und Vertrauenswürdigkeit auf der Ebene der Proben sowie auf der Ebene der Analyse charakterisiert. Die Charakterisierung resultiert im Programm in der Zuweisung eines Fehler- oder Statuswertes. Zu diesem Zeitpunkt in der Auswertung sind die Amplifikationskurven der selektierten Proben berechnet und korrigiert. Jede Probe besitzt einen Statuswert, der angibt, ob die berechnete Amplifikationskurve einen gültigen exponentiellen Bereich besitzt.

Zunächst wird versucht, anhand der selektierten Standards und der Angaben über den Schwellenwert die Kalibriergerade zu berechnen. Im Rahmen dieser Berechnungen wird eine Reihe von Tests zur Eignung der selektierten Standards durchgeführt, die in **Abb. 3-29** schematisch dargestellt ist.

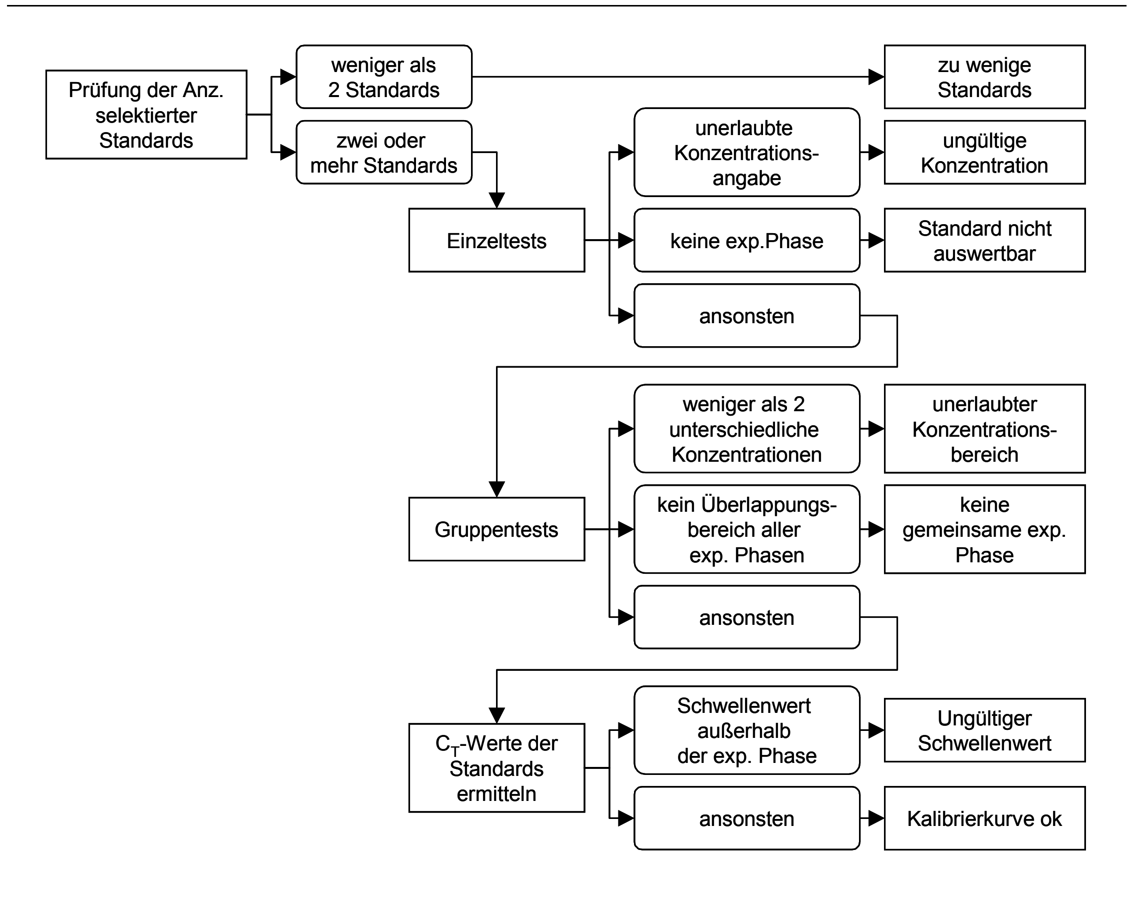


Abb. 3-29 Entscheidungsdiagramm zur Bestimmung des Fehlerwerts der quantitativen Analyse. Erklärung siehe Text.

Wenn mehr als zwei Standards selektiert sind, werden zunächst die Standards einzeln auf ihre Eignung geprüft. Ein Standard ist dann geeignet, wenn seine definierte Konzentration größer ist als Null und die Amplifikationskurve eine erkennbare exponentielle Phase besitzt. Sind alle selektierten Standards nach diesen Kriterien verwendbar, wird überprüft, ob die Standards mindestens zwei unterschiedliche Konzentrationen repräsentieren und ein gemeinsamer Überlappungsbereich ihrer bestimmten exponentiellen Phasen existiert. Ist dies der Fall, werden die C_T-Werte der Standards berechnet und in der Folge auch die Kalibriergerade ermittelt, sofern für jeden Standard auch ein C_T-Wert berechnet werden konnte. Wenn die automatische Bestimmung des Schwellenwerts ausgeschaltet ist, kann es vorkommen, dass der aktuelle Schwellenwert außerhalb der bestimmten exponentiellen Phase eines Standards liegt und damit die korrekte Berechnung eines C_T-Wertes für diese Probe nicht möglich ist. Nach der Bestimmung der Parameter der Kalibriergeraden werden schließlich noch die apparenten Amplifikationseffizienzen der Standards anhand ihrer exponentiellen Phasen ermittelt. Ist der Unterschied zwischen der minimalen und maximalen berechneten Effizienz der selektierten Standards größer als 0.3, wird ein Statuswert zurückgegeben, welcher vor möglichen Unstimmigkeiten wie z. B. verhältnismäßig schlecht amplifizierten Standards warnt.

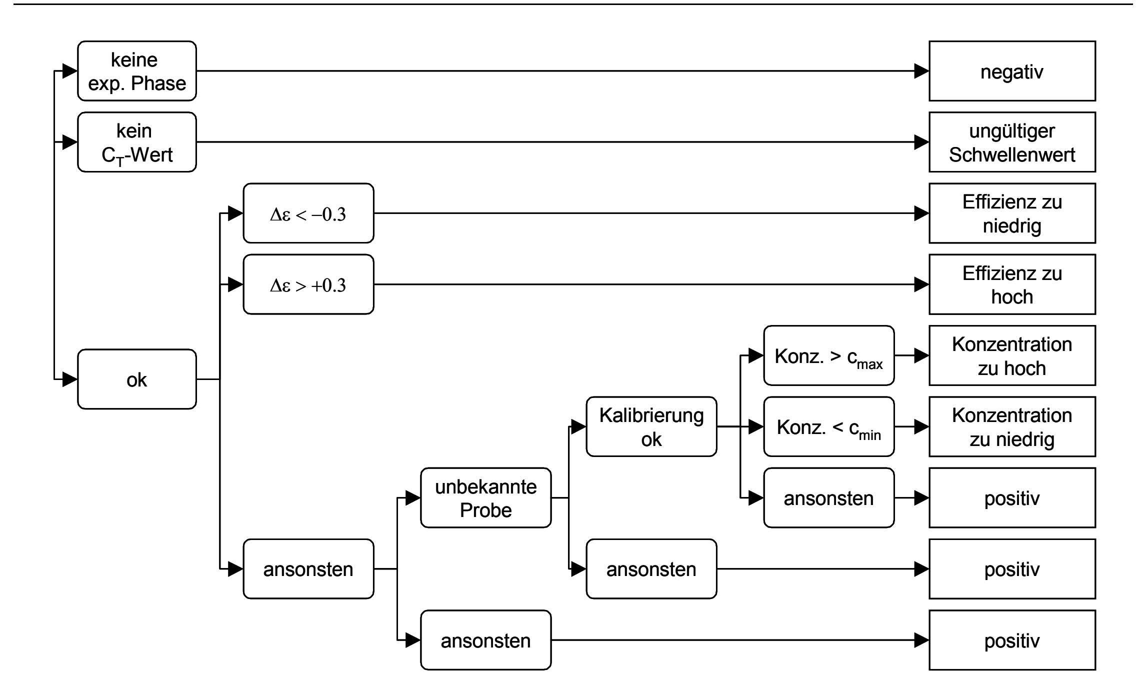


Abb. 3-30 Entscheidungsdiagramm zur Charakterisierung der Ergebnisse der Quantifizierung der selektierten Proben. $\Delta\varepsilon$ ist der Unterschied der berechneten Effizienz aus der exponentiellen Phase der jeweiligen Amplifikationskurve zur mittleren Effizienz aller selektierten Proben mit auswertbarer exponentieller Phase, c_{max} und c_{min} sind die größte bzw. kleinste Konzentration der selektierten Standards. Erklärung siehe Text.

Die abschließende Charakterisierung aller Proben im Einzelnen erfolgt im Anschluss an die Konzentrationsbestimmungen. Das Entscheidungsdiagramm dazu ist in **Abb. 3-30** dargestellt. Während der Berechnung der Amplifikationskurven aus den Rohdaten wurde bereits für jede Probe festgestellt, ob eine erkennbare bzw. auswertbare exponentielle Phase existiert. Konnte für die Probe keine exponentielle Phase ermittelt werden, wird sie als negativ eingestuft.

Besitzt die Amplifikationskurve eine exponentielle Phase, konnte aber für den aktuellen Schwellenwert kein C_T-Wert ermittelt werden, schneidet der Schwellenwert die Kurve nicht innerhalb ihrer exponentiellen Phase, was als ungültiger Schwellenwert bezeichnet wird. Besitzt die Probe eine erkennbare exponentielle Phase und konnte ein C_T-Wert berechnet werden, wird die Probe als positiv gewertet. Wenn es sich bei der betreffenden Probe um eine Unbekannte handelt und die Kalibrierung erfolgreich war, wird zusätzlich festgestellt, ob die berechnete Konzentration der Probe innerhalb des durch die Standards abgedeckten Konzentrationsbereichs liegt. Ist das nicht der Fall, wird ein entsprechender Statuswert zurückgegeben.

Die Charakterisierung der Quantifizierungsergebnisse für die Proben wird zur Anzeige des Status in den Ergebnissen in Abhängigkeit vom jeweiligen Probentyp interpretiert. Die Einstufung als „positiv" sollte z. B. im Falle einer Negativkontrolle mit einer Warnung gekenn-

zeichnet werden, wohingegen dieselbe Einstufung bei einer Positivkontrolle keine Warnung auslösen sollte. Die möglichen Ergebnisse mit den Symbolen zur Kennzeichnung der Einstufung sind in **Tab. 3-2** aufgeführt.

Tab. 3-2 **Interpretation der Charakterisierung und Symbole zur Darstellung.** Grüne Symbole zeigen an, dass bei der Charakterisierung und Interpretation der Ergebnisse dieser Probe keine Auffälligkeiten aufgetreten sind. Gelbe Symbole warnen vor möglicherweise fehlerhaften Ergebnissen, rote Symbole zeigen an, dass ein Standard nicht verwendet werden kann oder dass die Kontrolle als falsch-positiv oder falsch-negativ charakterisiert wurde.

Symbol	Bedeutung
Standards	
	Positiv: Gültiger Standard. Bei einer gültigen Kalibriergeraden ist auch die Konzentration dieses Standards anhand der Kalibriergerade sowie der Variationskoeffizient als prozentuale Abweichung der berechneten von der angegebenen Konzentration ermittelt.
	Ungültiger Schwellenwert. Der Schwellenwert befindet sich damit in jedem Falle außerhalb der Überlappungsbereichs der exponentiellen Phasen der selektierten Standards. Eine Kalibriergerade kann für diesen Schwellenwert nicht berechnet werden.
	Effizienz zu hoch. Das Ergebnis der gesamten Quantifizierung könnte fehlerhaft sein.
	Effizienz zu niedrig. Das Ergebnis der gesamten Quantifizierung könnte fehlerhaft sein.
	Die Amplifikationskurve des Standards hat keine erkennbare exponentielle Phase (negativ). Solange dieser Standard selektiert ist, kann keine Kalibriergerade berechnet werden.
Unbekannte	
	Positiv: Ein C_T-Wert ist berechnet, bei einer gültigen Kalibriergeraden auch die Konzentration dieser Probe.
	Negativ.
	Konzentration zu hoch: Die berechnete Konzentration ist nicht innerhalb des durch die selektierten Standards abgedeckten Konzentrationsbereichs.
	Konzentration zu niedrig: Die berechnete Konzentration ist nicht innerhalb des durch die selektierten Standards abgedeckten Konzentrationsbereichs.
	Ungültiger Schwellenwert: Es ist kein C_T-Wert berechnet, die Probe ist aber positiv, d. h. die Amplifikationskurve besitzt eine erkennbare exponentielle Phase.
	Effizienz zu hoch: Der ermittelte C_T-Wert und die bei gültiger Kalibriergerade berechnete Konzentration der Probe könnte falsch sein, weil beispielsweise ein falsches Produkt mit höherer Effizienz amplifiziert wurde.
	Effizienz zu niedrig: Der ermittelte C_T-Wert und die bei gültiger Kalibriergerade berechnete Konzentration der Probe könnte falsch sein, weil beispielsweise ein falsches Produkt mit geringerer Effizienz amplifiziert wurde.

Symbol	Bedeutung
Positivkontrollen	
⊕	Positiv: Ein C_T-Wert ist berechnet, bei einer gültigen Kalibriergeraden auch die Konzentration dieser Probe.
⧗	Ungültiger Schwellenwert: Es ist kein C_T-Wert berechnet, die Probe ist aber positiv, d. h. die Amplifikationskurve besitzt eine erkennbare exponentielle Phase.
ε	Effizienz zu hoch: Die Probe ist zwar positiv, jedoch ist möglicherweise ein falsches Produkt mit höherer Effizienz amplifiziert worden.
ε	Effizienz zu niedrig. Die Probe ist zwar positiv, jedoch ist möglicherweise ein falsches Produkt mit geringerer Effizienz amplifiziert worden.
⊖	Falsch negativ.
Negativkontrollen	
⊖	Negativ.
⧗	Ungültiger Schwellenwert: Es ist kein C_T-Wert berechnet, die Probe ist aber falsch-positiv, d. h., die Amplifikationskurve besitzt eine erkennbare exponentielle Phase.
ε	Effizienz zu hoch: Die Probe ist zwar falsch-positiv, jedoch ist möglicherweise ein falsches Produkt mit höherer Effizienz amplifiziert worden.
ε	Effizienz zu niedrig: Die Probe ist zwar falsch-positiv, jedoch ist möglicherweise ein falsches Produkt mit geringerer Effizienz amplifiziert worden.
⊕	Falsch positiv. Ein C_T-Wert ist berechnet, bei einer gültigen Kalibriergeraden auch die Konzentration dieser Probe.

Die vom Programm automatisch durchgeführten Charakterisierungen der Amplifikationskurven und deren Interpretation helfen dem Anwender, die erhaltenen Ergebnisse richtig zu werten. Zum einen wird dadurch sichergestellt, dass die Randbedingungen für eine quantitative Auswertung eingehalten werden, zum anderen werden Proben gekennzeichnet, die durch abweichende Amplifikationskurven auffallen. So ist z. B. die Konzentrationsbestimmung einer Probe mit einer Amplifikationskurve, deren exponentielle Phase nicht parallel (innerhalb gewisser Toleranzen) zu jenen der Standards verläuft, normalerweise nicht richtig. Die abweichende apparente Amplifikationseffizienz kann hier auf einen Fehler hinweisen. Möglicherweise wird bei dieser Probe eine andere Sequenz als die Zielsequenz (co-)amplifiziert oder die Zusammensetzung bzw. die Menge der Komponenten des Reaktionsansatzes sind fehlerhaft. Durch die entsprechende Warnung kann der Anwender auf solche möglichen Fehler aufmerksam gemacht werden und die Messung gegebenenfalls wiederholen.

3.2.4.7 Programmoberfläche

Für jedes geöffnete Experiment ist eine Vielzahl von Informationen anzuzeigen. Neben den Informationen zum Protokoll und den gemessenen Proben müssen verschiedene Datenreihen in Diagrammen grafisch dargestellt und die Ergebnisse von quantitativen Analysen und

Schmelzkurvenanalysen angezeigt werden. Die variablen Einstellungen zu den Analysen, wie z. B. das zu verwendende Datensegment und die auszuwertenden Kanäle, sollten jederzeit sichtbar sein und direkt geändert werden können, ebenso wie die Selektion der auszuwertenden Proben.

Eine rein parallele Darstellung aller Daten und Ergebnisse ist wegen des begrenzten Darstellungsbereichs auf dem Monitor nicht sinnvoll, ebenso wenig wie eine rein sequenzielle Darstellung, bei welcher immer nur ein Diagramm oder eine Tabelle exklusiv angezeigt wird, da die gleichzeitige Darstellung verschiedener Kombinationen von Diagrammen und Tabellen für eine gute Übersicht und Interpretation der Ergebnisse sinnvoll sein kann, was letztlich von der Fragestellung und dem Anwender abhängt.

Die hier gewählte Lösung besteht in der Darstellung der Diagramme und Tabellen in separaten Fenstern, die innerhalb des Anwendungsfensters beliebig positioniert werden können, in Kombination mit einer einfachen Navigation von den Einstellungen zu den zugehörigen Fenstern sowie der Möglichkeit, Gruppen von Fenstern (z. B. mit Anzeigen zur Schmelzkurvenanalyse) ein- und ausblenden zu können. Dazu wurde die Anwendung als *multiple document interface* (MDI) Applikation programmiert. Der Begriff MDI wird von Microsoft zur Bezeichnung eines Haupt- bzw. Anwendungsfensters (*main window* oder *MDI-parent window*) verwendet, welches mehrere sogenannte Dokumentfenster (*client windows* oder *MDI-child windows*) verwaltet und innerhalb seiner eigenen Oberfläche darstellen kann. Die Dokumentfenster dienen hier jedoch nicht der Darstellung getrennter Dokumente, sondern der Darstellung von Informationen, Messdaten und Ergebnissen von RT-PCR-Experimenten.

Die Auswahl der auszuwertenden Proben sowie benutzerdefinierte Einstellungen zur Auswertung werden im Hauptfenster vorgenommen, welches in **Abb. 3-31** gezeigt ist. Im Folgenden wird zunächst ein Überblick über die Elemente des Hauptfensters gegeben, woran sich eine kurze Beschreibung der Dokumentfenster im Einzelnen anschließt.

Die verfügbaren Proben werden sortiert angezeigt in einer Baumsstruktur mit den Ebenen Experiment, Probentyp, Gruppe (sofern vorhanden) und Proben. Die Selektion der auszuwertenden Proben erfolgt durch einen Klick auf die Auswahlfelder. Ein schwarzes Kreuz (☒) kennzeichnet die Probe oder Gruppe als ausgewählt, ein leeres Feld (☐) als entsprechend nicht selektiert. Gruppen mit sowohl selektierten als auch nicht selektierten Proben sind durch ein graues Kreuz (☒) gekennzeichnet. Die rechts daneben angezeigten Attribute zur grafischen Darstellung der Messwerte können durch einen Doppelklick auf die Symbole geändert werden. Grundsätzlich werden in den Dokumentfenstern nur Daten und Ergebnisse selektierter Proben angezeigt.

Die Sektionstafeln zeigen in ihren Titelzeilen links das Symbol eines Ordners in entweder geöffneter ( ) oder geschlossener ( ) Darstellung. Bei einem geschlossenen Ordner werden alle zu der jeweiligen Sektion gehörenden Eingabefelder und Dokumentfenster nicht angezeigt. Dadurch kann die Übersicht erhöht werden, indem nicht interessierende Teile der Darstellung vollständig ausgeblendet werden. Durch einen Klick auf das Ordnersymbol lässt sich zwischen den Zuständen offen und geschlossen wechseln. Auf der rechten Seite in den Titelzeilen der Sektionstafeln befinden sich Fenstersymbole, welche die zur Sek-

tion gehörenden Dokumentfenster symbolisieren (, , , ,). Ein Klick auf ein Fenstersymbol öffnet das entsprechende Dokumentfenster und bringt es in den Vordergrund. Damit wird eine sehr übersichtliche und direkte Ansteuerung der unterschiedlichen Anzeigen möglich.

Im Folgenden werden die Elemente der vier Sektionen mit den zugehörigen Dokumentfenstern im Einzelnen kurz beschrieben.

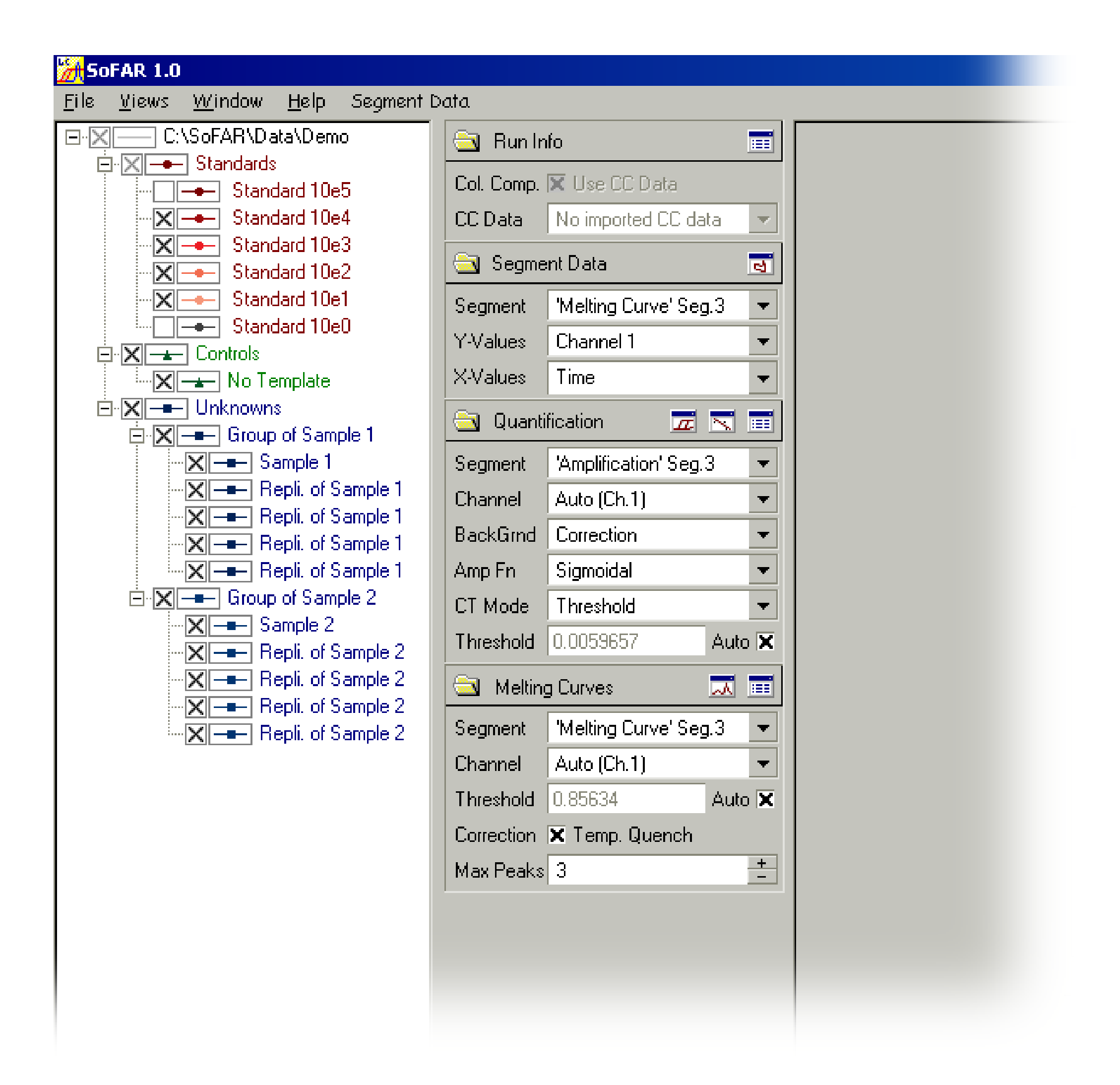


Abb. 3-31 Anwendungsfenster von SoFAR. Auf der linken Seite des Anwendungsfensters befindet sich die Komponente zur Selektion der Proben, eine Instanz von `TLCTreeview` (siehe Tab. 7-2 und Abb. 7-1). Gleich daneben befindet sich eine Leiste mit den Feldern zur Anzeige und Eingabe der Einstellungen zur Datenanalyse, die sich in vier Sektionen (Sektionstafeln) gliedert: *„Run Info“*, *„Segment Data“*, *„Quantification“* und *„Melting Curves“*, welche zum einen die Eingabefelder zusammenfassen und zum anderen Schaltflächen zur Anzeige der Dokumentfenster enthalten.

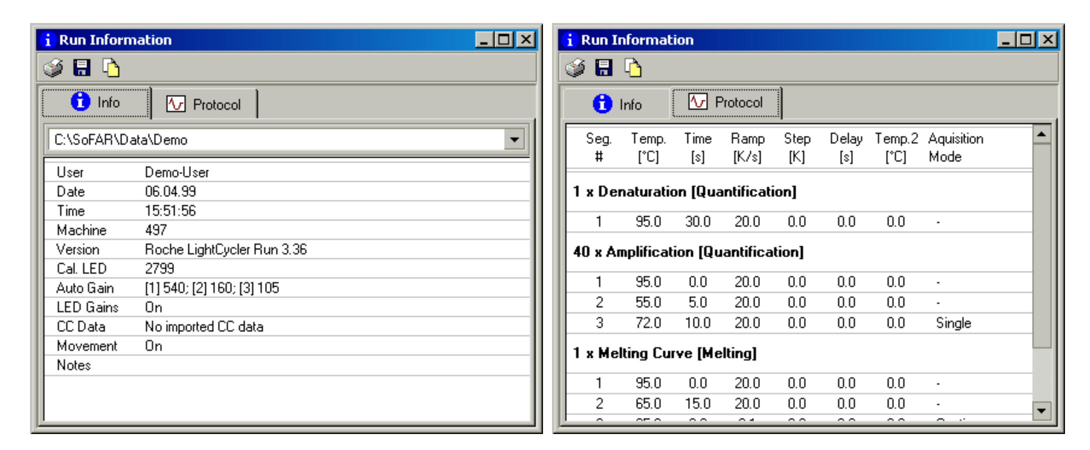


Abb. 3-32 Dokumentfenster der Sektion „*Run Info*". Gezeigt ist das Fenster einmal mit dem selektierten Tabellenblatt „*Info*" und einmal mit dem selektierten Tabellenblatt „*Protocol*". Im Tabellenblatt „*Info*" befindet sich ein Auswahlfeld zur Selektion des geöffneten Experimentes, dessen Daten in der Tabelle angezeigt werden (im gezeigten Beispiel ist das Experiment „C:\SoFAR\Data\Demo" ausgewählt).

● ***Run Info***. Die Sektion „*Run Info*" dient der Anzeige allgemeiner Informationen zu den geöffneten Experimenten wie Benutzer, Datum, Geräteeinstellungen und dem PCR-Protokoll, welches das Temperaturprofil und die Messpunkte festlegt. Außerdem werden importierte Daten mit Informationen zur Farbkompensation bei Multiplexansätzen angezeigt („CC *Data*"), die hier auch selektiert werden können. Die Liste mit Farbkompensationsdaten umfasst alle importierten Daten aller geöffneten Experimentdateien. Die Selektion eines Eintrages aus dieser Liste bewirkt die Anwendung dieser Daten zur Farbkompensation für alle Messwerte aller geöffneten Dateien, sofern angegeben ist, dass eine Farbkompensation durchgeführt werden soll („*Use CC Data*").

Die Sektion enthält ein Dokumentfenster (, „*Run Information*") zur Anzeige der Protokolldaten und Geräteeinstellungen (**Abb. 3-32**). Die Angaben über z. B. Benutzer, Datum und Zeit müssen bei den gleichzeitig geöffneten Experimenten nicht identisch sein. Daher lässt sich in einem Auswahlfeld das geöffnete Experiment auswählen, dessen Angaben in der darunter befindlichen Tabelle angezeigt werden.

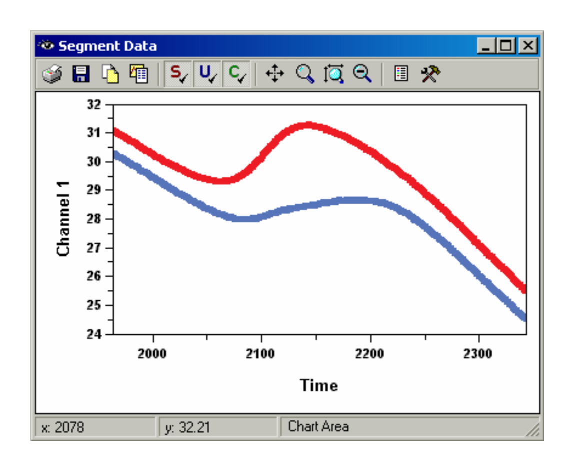


Abb. 3-33 Dokumentfenster der Sektion „*Segment Data*". Das Fenster dient der Anzeige von Messwerten der selektierten Proben in einem x/y-Diagramm.

● ***Segment Data***. In der Sektion „*Segment Data*“ werden die Datenreihen ohne weitere Analyse dargestellt. Die Eingabefelder ermöglichen die Auswahl des Datensegments („*Segment*“) sowie die Festlegung der Datenquelle für die x- und y-Werte im Diagramm des Dokumentfensters („*X-Values*“ bzw. „*Y-Values*“). Diese Sektion enthält ebenfalls nur ein Dokumentfenster (, „*Segment Data*“), welches die gewählten Messwerte der selektierten Proben in Form eines x/y-Diagramms anzeigt (**Abb. 3-33**).

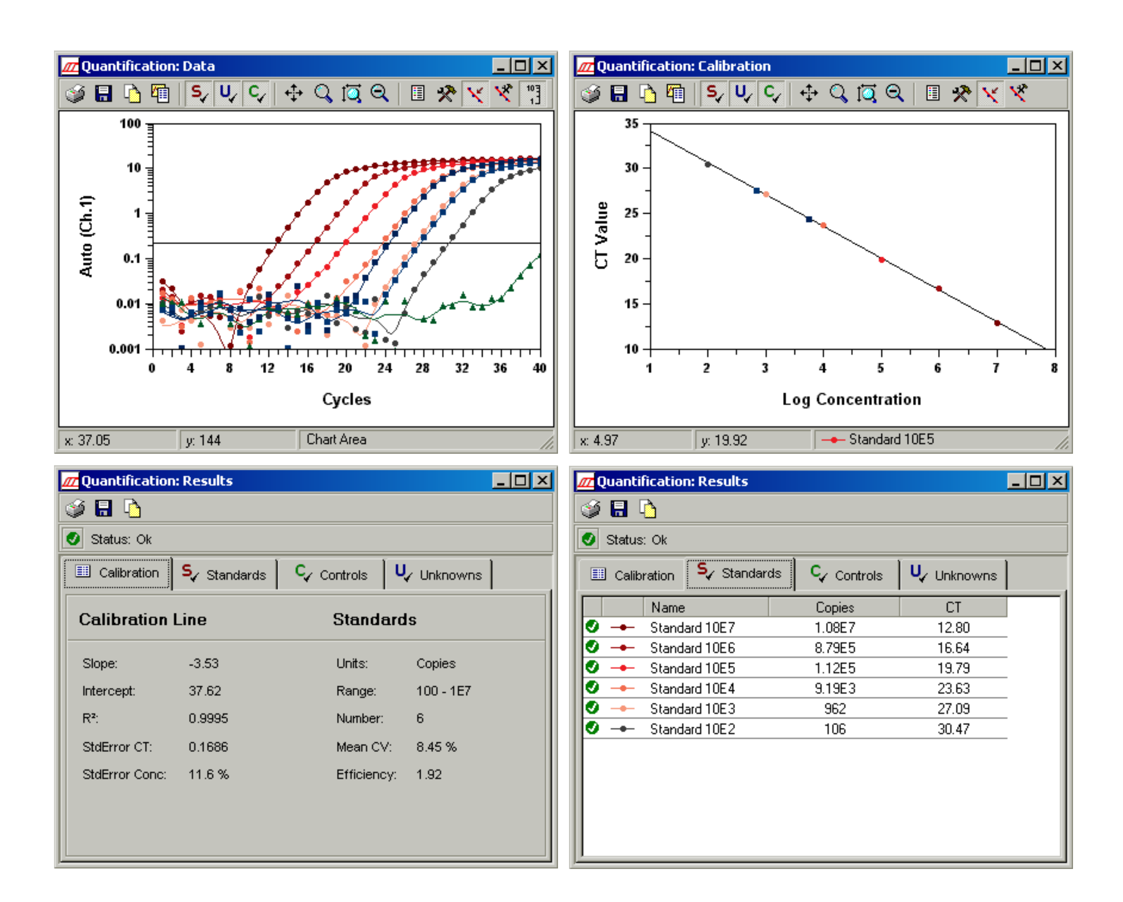


Abb. 3-34 Dokumentfenster der Sektion „*Quantification*“. Das linke obere Fenster („*Data*“) dient der Darstellung der Amplifikationskurven. Die horizontale Linie stellt den aktuell gewählten Schwellenwert dar. Wenn die automatische Bestimmung des Schwellenwerts ausgestellt ist, kann der Schwellenwert durch Klicken und Ziehen der Linie mit der Maus direkt im Diagramm geändert werden. Das rechte obere Fenster („*Calibration*“) dient der Darstellung der Kalibriergeraden, wo außer den C_T-Werten der selektierten Standards auch die der selektierten Kontrollen und Unbekannten angezeigt werden können. Die unteren beiden Abbildungen zeigen das Ergebnisfenster („*Results*“) links mit „*Calibration*“, rechts mit „*Standards*“ als selektiertem Blatt. Die Blätter „*Controls*“ und „*Unknowns*“ entsprechen in ihrem Aufbau dem Blatt „*Standards*“.

● ***Quantification***. Die Sektion „*Quantification*“ fasst die zur Quantifizierung wichtigen Eingabefelder zusammen. In der Liste der Datensegmente („*Segment*“) sind hier nur diejenigen Datensegmente aufgeführt, die Teil eines PCR-Programms sind, für welches der Ana-

lysemodus „*Quantification*“ gesetzt ist. Im Eingabefeld „*Channel*“ kann der Kanal bzw. das Kanalverhältnis bestimmt werden, von welchem die Fluoreszenzwerte zur Auswertung verwendet werden sollen. Weiterhin befinden sich in der Sektionstafel Listenfelder zur Auswahl des Verfahrens zur Hintergrundkorrektur („*BackGrnd*“), der Amplifikationsfunktion zur Anpassung an die exponentiellen Phasen („*Amp Fn*“) und der Methode zur Bestimmung der C_T-Werte („*CT Mode*“). Je nach Wahl der Methode zur Bestimmung der C_T-Werte („*Threshold*“ für die in dieser Arbeit entwickelte Schwellenwert-Methode oder „*Noiseband*“ für die *Fit-points*-Methode) werden Eingabefelder angezeigt, über welche die Lage des Schwellenwerts manuell bestimmt und die Option der automatischen Bestimmung des Schwellenwerts eingestellt werden kann. Über ein Auswahlfeld („*Auto*“) kann definiert werden, ob der Schwellenwert automatisch eingestellt werden soll.

Daten und Ergebnisse zur Quantifizierung werden in insgesamt drei Dokumentfenstern dargestellt (**Abb. 3-34**): einem Diagrammfenster zur Anzeige der Amplifikationskurven (, „*Data*“), ein weiteres zur Darstellung der Kalibriergeraden (, „*Calibration*“) sowie ein Ergebnisfenster zur Auflistung der Ergebnisse auf mehreren Tabellenblättern (, „*Results*“). Die Tabellenblätter umfassen ein Blatt zur Anzeige der Ergebnisse der Kalibriergeraden mit verschiedenen statistischen Kennwerten wie Korrelationskoeffizienten und Standardfehler sowie die aus der Steigung der Standardgeraden berechnete Effizienz. Die übrigen drei Blätter enthalten je Probentyp (Standards, Kontrollen und Unbekannte) eine Tabelle mit den Einzelergebnissen für die berechneten C_T-Werte, Konzentrationen und Charakterisierungen.

- ***Melting Curves***. Diese Sektion beinhaltet die zur Schmelzkurvenanalyse wichtigen Eingabefelder und Dokumentfenster. Wie in der Sektion „*Quantification*“ können auch hier das Segment mit den auszuwertenden Daten sowie der Kanal bzw. das Kanalverhältnis

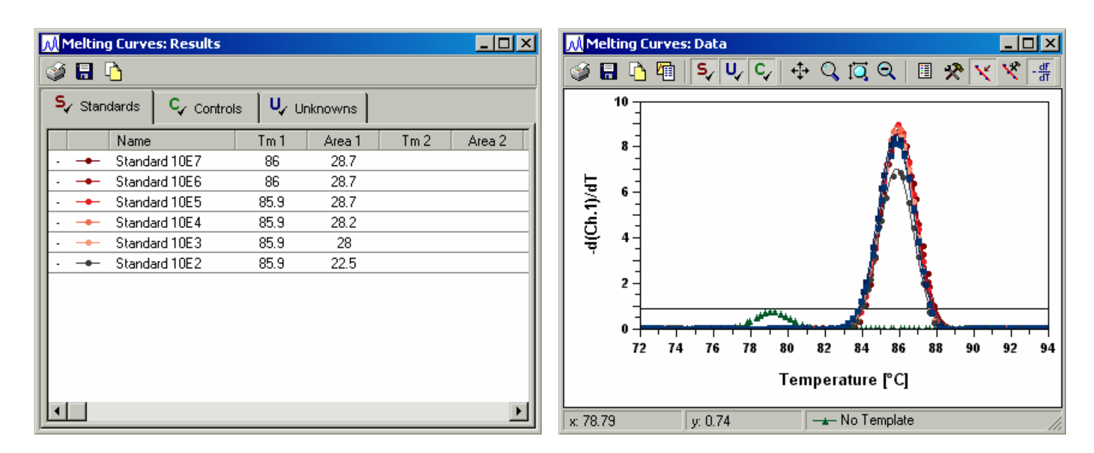


Abb. 3-35 Dokumentfenster der Sektion „*Melting Curves*“. Das links gezeigte Fenster („*Data*“) stellt die Schmelzkurvendaten bzw. deren Peakkurven in einem x/y-Diagramm dar. Die horizontale Linie in der gezeigten Peakkurvenansicht zeigt den Schwellenwert zur Peakdetektion an. Im rechten Fenster („*Results*“) sind die Ergebnisse der Schmelzkurvenanalyse dargestellt, ebenfalls sortiert nach den Probentypen, die in separaten Tabellenblättern dargestellt werden („*Standards*“, „*Controls*“ und „*Unknowns*“).

gewählt werden, dessen Fluoreszenzwerte zur Auswertung verwendet werden sollen („*Segment*“ und „*Channel*“). Hier werden im Listenfeld „Segment“ jedoch nur diejenigen Datensegmente aufgeführt, welche Teil eines Programms mit dem Analysemodus „*Melting Curve*“ sind. Weiterhin kann der Schwellenwert zur Peakdetektion gewählt werden („*Threshold*“) oder seine automatische Wahl bestimmt werden („*Auto*“). Über ein Auswahlfeld („*Correction*“) lässt sich angeben, ob die Schmelzkurven um den temperaturbedingten Quench korrigiert werden sollen. Schließlich kann noch die Anzahl maximal zu detektierender Peaks pro Kurve eingestellt werden („*Max* Peaks“). Die Sektion beinhaltet zwei Dokumentfenster zur Anzeige der Datenreihen sowie der Ergebnisse (**Abb. 3-35**. Die im Diagrammfenster (, „*Data*“) dargestellten Datenreihen können entweder in Form der Schmelzkurven oder der Peakkurven in einem x/y-Diagramm angezeigt werden. Die in den Tabellenblättern des Ergebnisfensters (, „*Results*“) aufgeführten Ergebnisse beschreiben die ermittelten Schmelztemperaturen und Peakflächenintegrale der Proben, ebenfalls nach Probentyp sortiert.

3.3 Anwendungsbeispiele in der molekularen Tumordiagnostik

Immer effektivere und spezifischere Untersuchungsmethoden liefern einerseits immer schlüssigere diagnostische Informationen, sind aber andererseits - jedenfalls soweit invasiv - mit Risiken behaftet und verführen durch die Attraktion kraft des Neuen oder durch den Zwang zur Amortisation zum Einsatz ohne strenge Indikation.

Hans-Peter Wolff

In den folgenden drei Abschnitten werden Beispiele zur Quantifizierung wichtiger, tumorrelevanter genetischer Elemente mit den Mitteln der RT-PCR mit externer Standardisierung dargestellt. Die Beispiele demonstrieren die Anwendbarkeit dieser Technik auf hochamplifizierte mobile genetische Elemente, deren Anzahl im Genom großen Schwankungen unterliegt, auf mäßig amplifizierte Onkogene sowie auf deletierte Tumorsuppressorgene, die hohe Anforderungen an die Genauigkeit der Technik stellen.

3.3.1 Bestimmung der Kopienzahl transponierter Elemente

Der Platyfisch (*Xiphophorus maculatus*) ist eines der ältesten Tiermodelle zum Studium der Tumorentstehung. Bereits 1928 entdeckten M. GORDON, G. HÄUSSLER und K. KOSSWIG unabhängig voneinander, dass bestimmte F1-Hybride von *Xiphophorus maculatus* und *Xiphophorus helleri* (Schwertträger) Melanome entwickeln.[221-223] Die Untersuchungen dieser nach Mendel vererbten Anlage zur Tumorentwicklung bildeten die Grundlage zur Etablierung des *Xiphophorus*-Systems als Modell für die Entwicklung und Vererbung humaner Melanome und anderer Krebsarten, das wesentlich von F. und A. ANDERS entwickelt wurde.[224-]

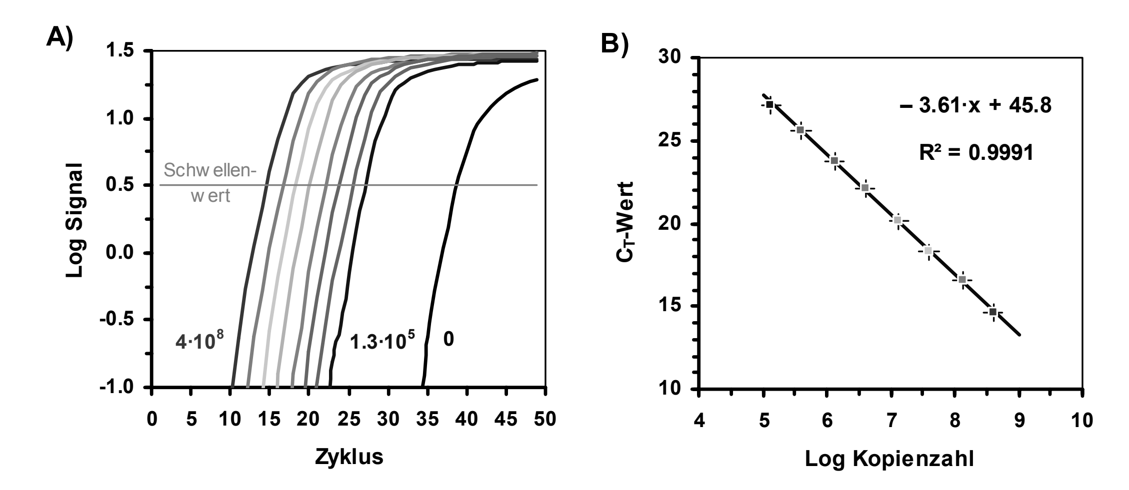


Abb. 3-36 XIR-2.5-Standardreihe. A) Amplifikationskurven der Standards (Kanal 1). Die Lage des Schwellenwertes zur Bestimmung der C_T-Werte ist eingezeichnet. Die Probe „0" enthält kein Templat. **B)** Kalibrierkurve aus den C_T-Werten der Standards aus A). Die Gleichung der Regressionsgeraden sowie das Bestimmtheitsmaß sind im Diagramm angegeben.

[226] Hieran wurden in den 1960er Jahren erstmals Onkogene und Tumorsuppressorgene nachgewiesen.[149]

Im Rahmen dieser Forschungen kam dabei die Frage auf, welche Rolle die Aktivität eines Retrotransposons bei der Entstehung bestimmter erblicher, aber nicht nach Mendel vererbter Tumoren spielt.[227] Retrotransposons sind „mobile" genetische Elemente, welche kopiert und an anderen Stellen im Genom integriert werden können. Der ungerichtete Einbau eines transponierbaren Elements in genomische Sequenzen, welche für die Regulation des Zellzyklus von Bedeutung sind, kann eine Transformation der Zelle zu einer Tumorzelle begünstigen, indem z. B. Tumorsuppressorgene ausgeschaltet werden.[203] Durch die Anwendung der RT-PCR mit SYBR-Green I als Fluorophor war es möglich, die genomische Kopienzahl des ursprünglich in *X. helleri* nachgewiesenen Retrotransposons XIR-2.5 für verschiedene Stämme, aber auch für zahlreiche Wildarten der Gattung *Xiphophorus* und weitere Gattungen der Familie *Poeciliidae* schnell und präzise zu bestimmen.

Die Quantifizierung der XIR-2.5-Kopienzahl erfolgte anhand einer seriellen 1:3 Verdünnungsreihe des linearisierten Plasmids pXIR-2.5 als Standard. Die Kopienzahlen der Standards in den Reaktionsansätzen betrugen zwischen $4 \cdot 10^8$ und $1.33 \cdot 10^5$, der untersuchte Konzentrationsbereich überdeckt damit mehr als 3 Größenordnungen (**Abb. 3-36**). Sowohl Schmelzkurvenanalysen als auch PAGE-Analysen zeigen, dass bei allen untersuchten Arten sowie bei den Standards dasselbe spezifische 223 Bp lange PCR-Produkt entsteht.

Die Ergebnisse der Bestimmung der genomischen Kopienzahlen von XIR-2.5 für verschiedene Arten, Kreuzungen und Zuchtstämme aus der Familie *Poeciliidae* sind in **Abb. 3-37** gezeigt. Die Standardabweichungen von Fünffach-Bestimmungen pro Fisch liegen unter 10 %, während die Variationen von Fisch zu Fisch teilweise sehr groß sind (siehe *X.*

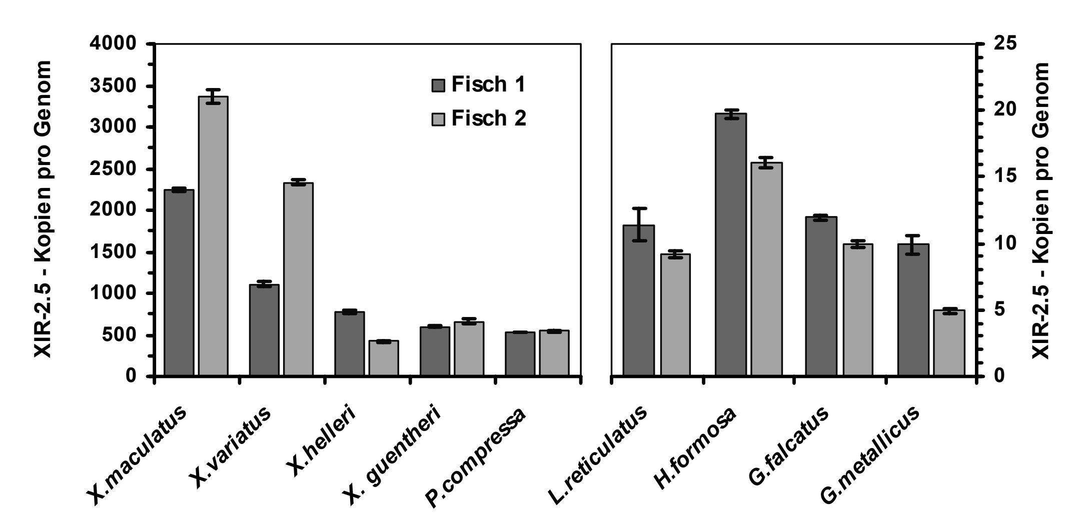


Abb. 3-37 XIR-2.5 Kopien pro Genom in verschiedenen Arten der Familie *Poeciliidae*. Es wurden je zwei Fische jeder Art untersucht. Die Ergebnisse sind die Mittelwerte von Fünffach-Bestimmungen, die Fehlerbalken entsprechen den einfachen Standardabweichungen. Die Ergebnisse für Arten, für die genomische XIR-2.5-Kopienzahlen von weniger als 250 bestimmt wurden, sind im rechten Diagramm mit einer abweichend skalierten Größenachse gezeigt.

maculatus und *X. variatus* in **Abb. 3-37**), die nicht durch Messfehler zu erklären sind. Um Aussagen über die tatsächliche individuelle Variation der genomischen XIR-2.5-Kopienzahl treffen zu können, muss eine deutlich größere Zahl von Individuen jeder Art untersucht werden.

Die Arten lassen sich offensichtlich in zwei Gruppen mit deutlich unterschiedlichen genomischen XIR-2.5-Kopienzahlen einteilen. Die eine Gruppe mit den Arten der Gattungen *Xiphophorus* und *Priapella* sowie Hybride dieser Arten weisen zwischen 500 und 1000 XIR-2.5-Kopien je Genom auf, wobei die beiden eng verwandten Arten *X. maculatus* und *X. variatus* sogar 2000 bis 3000 Kopien erreichen (großes Diagramm in **Abb. 3-37**). Die andere Gruppe bilden die Arten der Gattungen *Girardinus*, *Heterandria* und *Lebistes* (Guppy) mit Kopienzahlen zwischen 5 und 20 (rechtes Diagramm in **Abb. 3-37**). Die klare Unterscheidbarkeit dieser Gruppen anhand ihrer XIR-2.5-Kopienzahlen kann zur systematischen Einordnung der phänotypisch mitunter schwer unterscheidbaren Arten herangezogen werden.

3.3.2 Bestimmung des Amplifikationsgrades von Onkogenen

Onkogenamplifikationen finden sich in vielen Tumoren und sind meistens verbunden mit einer schlechten Prognose für den Krankheitsverlauf. Tumore mit amplifizierten Onkogenen sind in der Regel aggressiver und resistenter gegen therapeutische Maßnahmen.[137] Die Diagnose des Status bestimmter Onkogene ist daher von Bedeutung für die Wahl einer adä-

quaten Therapieform für den Patienten. Die Bestimmung des Ausmaßes der Amplifikation kann auf drei Ebenen erfolgen: (i) auf der Proteinebene durch immunohistochemischen Verfahren, (ii) auf der Transkriptebene durch quantitative Reverse-Transkriptase-PCR-Techniken sowie (iii) auf der genomischen Ebene z. B. durch FISH- oder quantitative PCR-Techniken.

Das hier entwickelte Verfahren erlaubt die Quantifizierung der onkogenen Amplifikation von *HER-2/neu*, die in 20-30 % der Fälle von Brustkrebs vorliegt [136] und mit einer höheren Rückfallwahrscheinlichkeit nach einer Therapie sowie einer verkürzen Lebenserwartung korreliert.[137] Die Heilungs- und Überlebenschancen können in den Fällen, in denen tatsächlich eine Überexpression von HER-2 vorliegt, durch die Applikation eines gegen den Rezeptor gerichteten monoklonalen Antikörpers (Herceptin) deutlich verbessert werden.[138,139] Daher ist ein einfach und schnell anwendbares, robustes Verfahren zur Diagnose einer HER-2-Überexpression von großer klinischer Bedeutung.

Die Quantifizierung der *HER-2/neu*-Kopienzahlen (N_T, T:Testgen) werden zur Berechnung des Amplifikationsstatus auf die Kopienzahlen des *IGF-1*-Promotors (N_R, R:Referenzgen) bezogen, der hier als Referenzsequenz zur Bestimmung der Zahl der Genomäquivalente in der Probe dient: $q = N_T : N_R$. Die Bestimmung von q erfolgt für die Test- bzw. Tumorprobe ($q_T = (N_T)_T : (N_R)_T$) sowie für eine Referenz- bzw. Normalprobe ($q_N = (N_T)_N : (N_R)_N$) derselben Person. Der normierte Amplifikationsstatus Q für das Testgen berechnet sich daraus mit

$$Q = \frac{q_T}{q_N} = \frac{(N_T)_T \cdot (N_R)_N}{(N_R)_T \cdot (N_T)_N} \quad \text{Gl. 3.22}$$

IGF-1 wurde hier erstmals als Referenzgen gewählt, weil der Genlokus auf dem langen Arm des Chromosoms 12 in Mammakarzinomen nur selten selbst von Veränderungen betroffen ist: Die Chromosomen 2, 6, und 12 sind die am seltensten in ihrer Anzahl veränderten, und nur vier von insgesamt 1638 Untersuchungen berichten über strukturelle Veränderungen der *IGF-1*-Promotorregion (12q22).[228]

Zur Amplifikation der Zielsequenzen werden die Primerpaare HER-2-H2/R2 und IGF-H1/R1 verwendet. Die Signalgenerierung während der PCR erfolgt für beide Zielsequenzen durch Hybridisierungssonden, wobei das Signalverhältnis von Akzeptor- zu Donorfluorophor ausgewertet wurde. Damit kann dieses Verfahren ohne großen Aufwand zu einem Multiplex-Verfahren modifiziert werden, bei dem beide Zielsequenzen parallel im selben Reaktionsansatz quantifiziert werden. Dazu muss lediglich der Akzeptorfluorophor einer der beiden Akzeptorsonden durch RED-705 ersetzt werden, der im dritten Kanal des LightCyclers detektiert werden kann.

Die Präzision und Linearität der Bestimmung von Q wurde in Simulationsexperimenten ermittelt (**Abb. 3-38**). Dazu wurden die Quantifizierungsergebnisse unterschiedlich konzentrierter Proben mit genomischer DNA von Blutlymphozyten gesunder Personen verwendet. Um z. B. eine zehnfache Amplifikation des Testgens (*HER-2/neu*) gegenüber dem Referenzgen (*IGF-1*) zu simulieren ($Q = 10$), wurde die *HER-2/neu*-Kopienzahl in der unver-

dünnten Probe gemessen, während die *IGF-1*-Kopienzahl in der 1:10-verdünnten Probe bestimmt wurde. Die DNA-Menge im Reaktionsansatz zur Messung der *IGF-1*-Kopienzahl betrug dabei immer 1.5 ng/µl, was etwa 500 Genomäquivalenten entspricht.

Mit dem vorgestellten Verfahren lassen sich die Q-Werte für eine Deletion ($Q = 0.5$) ebenso wie für eine zehnfache Amplifikation ($Q = 10$) mit hoher Präzision messen (**Abb. 3-38**). Die Variationskoeffizienten der Fünffach-Bestimmungen liegen bei ca. 6 % (für die Deletion bei 11 %). Die Anhand von *pair-wise randomization tests*[209] ermittelten p-Werte für die Messwertgruppenpaare von $Q = 0.5/1.0$ und $Q = 5/10$ liegen jeweils bei 1.0, was bedeutet, dass innerhalb des Messbereichs Unterschiede von einem Faktor 2 hochsignifikant unterschieden werden können. Der *pair-wise randomization test* wird angewendet, um die Wahrscheinlichkeit zu ermitteln, mit welcher zwei Verteilungen unterschiedlicher Stichproben nicht Teile derselben Grundgesamtheit repräsentieren. Diese Methode lässt sich auf alle Verteilungen anwenden, da sie keine Annahme über die Art der Verteilungen macht, während der üblicherweise verwendete t-Test (oder *Students-Test*) nur für Normalverteilungen sinnvolle Ergebnisse liefert.

Die Genauigkeit der Quantifizierung wurde anhand von Zelllinien mit bekanntem *HER-2/neu*-Amplifikationsstatus demonstriert (**Abb. 3-39**). Die Standardreihen für beide Zielsequenzen, erstellt anhand serieller 1:2-Verdünnungsreihen humaner genomischer DNA, decken einen Konzentrationsbereich von 125 bis 8000 Genomäquivalenten je Reaktionsansatz ab. Die aus den Steigungen der Kalibriergeraden berechneten mittleren Amplifikationseffizienzen sind für *HER-2/neu* mit 1.85 etwas geringer als für *IGF-1* mit 1.94 (**Abb. 3-39 B**). Als Referenzprobe zur Bestimmung des q_N-Werts wurde eine Blutlymphozyten-DNA verwendet (Kontrolle).

Die Standardabweichungen der Fünffach-Bestimmungen von N_T und N_R liegen jeweils bei ca. 5 %, woraus sich für die Standardabweichungen von Q Werte zwischen 5 % und 12 % (MW: 7.5 %) ergeben. Die Ergebnisse in den untersuchten Zelllinien stimmen sehr gut mit den in der Literatur beschriebenen Werten überein,[229-232] was die Eignung dieses Tests belegt.

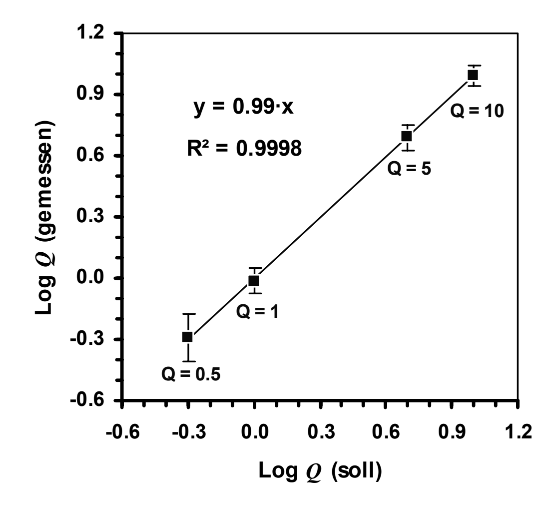


Abb. 3-38 Ergebnis des Simulationsexperiments zur *HER-2/neu*-Genamplifikation. Die Durchführung ist im Text beschrieben. Aufgetragen sind die experimentell bestimmten Q-Werte (Mittelwerte mit Standardabweichung von 5fach-Bestimmungen) gegen die Sollwerte für Q, die sich aus den bekannten Konzentrationsverhältnissen ergeben. Die Sollwerte von Q sind jeweils unter den Datenpunkten angegeben.

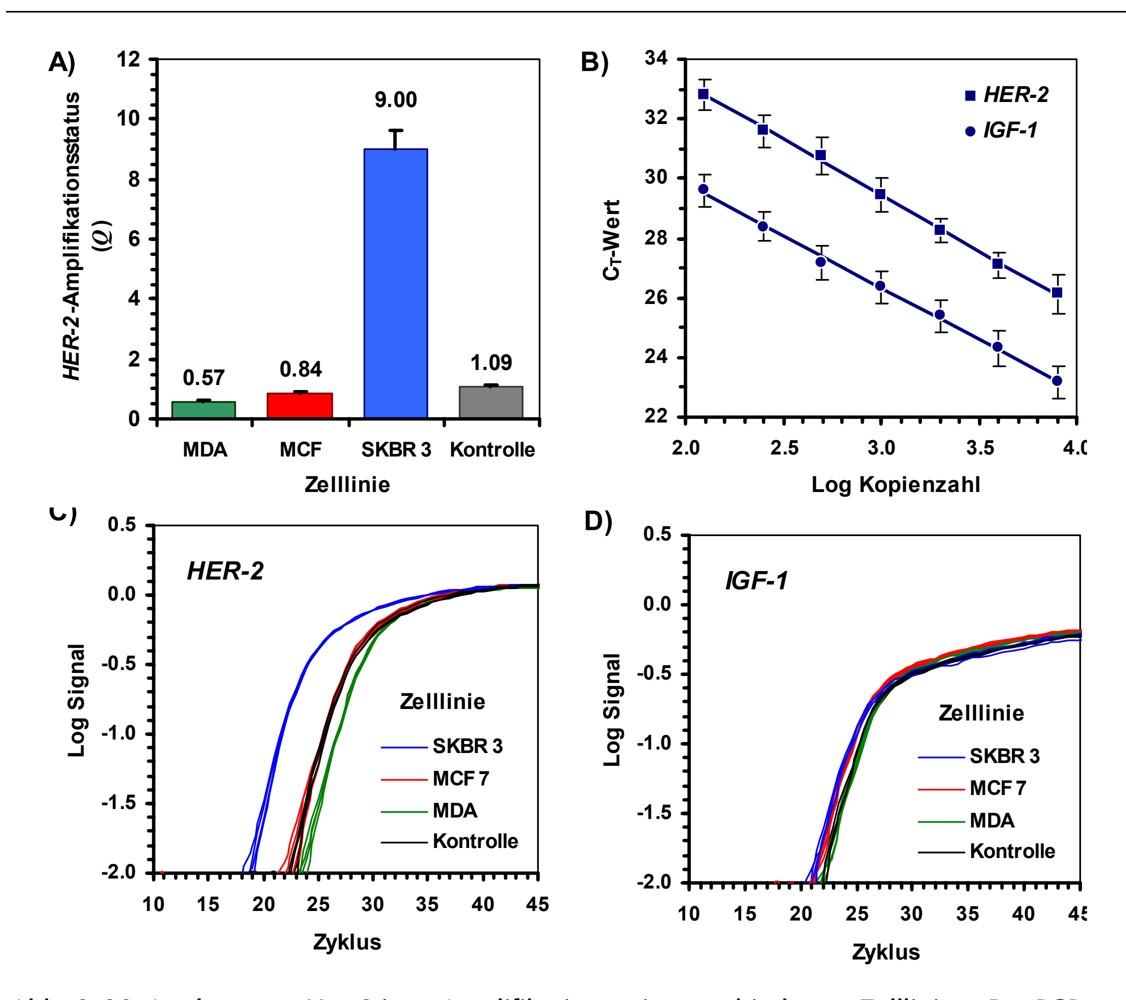


Abb. 3-39 Analyse von *Her-2/neu* Amplifikationen in verschiedenen Zelllinien. Die PCR erfolgte unter Standardbedingungen mit den Primerpaaren HER-H1/R1 und IGF-H1/R1, die Detektion mit den sequenzspezifischen Hybridisierungssonden. Ausgewertet wurde das Signalverhältnis von Akzeptor zu Donor (Kanal 2/Kanal 1). Als Probentemplat wurden jeweils 4 ng genomische DNA eingesetzt. MDA: MDA-468, MCF: MCF-7, SKBR3: SK-BR-3, Kontrolle: Blutlymphozyten-DNA einer gesunden Person. **A)** Ergebnis der Bestimmung des *Her-2/neu*-Amplifikationsstatus in Zelllinien. **B)** Mittlere Kalibriergeraden für beide Zielsequenzen von sechs Einzelbestimmungen aus verschiedenen Experimenten (Mittlere C_T-Werte und Standardabweichung). **C)** und **D)** Amplifikationskurven der untersuchten Proben (jeweils fünffach) für beide Zielsequenzen.

Abschließend ist das Ergebnis der Untersuchung einer realen Tumorprobe gezeigt, die freundlicherweise von M. KÖNIGSHOFF zur Verfügung gestellt wurde. Als Referenzprobe diente hier die Blutlymphozyten-DNA des Patienten. Die Tumor-DNA stammt aus mikrodissektiertem Gewebe eines großen intraduktalen Karzinoms einer Patientin mit positivem Östrogen- und Progesteron-Hormonrezeptorstatus und wurde wie in **2.4.1** beschrieben isoliert. Zum Vergleich wurde diese Probe anhand zweier dem Genlokus benachbarter polymorpher Marker auch hinsichtlich ihrer *HER-2/neu*-Allelverteilung untersucht. Als Marker dazu dienten die beiden am stärksten mit *HER-2/neu* koppelnden $d(AC)_n$-Dinukleotid-STRs

D17S800 und D17S1818, die sich 0.9 cM proximal bzw. 0.5 cM distal vom Genlokus befinden. Das Verfahren entspricht einer LOH-Analyse (siehe auch **3.3.3**), nur dass hier die gefundene Allelungleichverteilung nicht als Heterozygositätsverlust aufgrund der Deletion des einen Allels, sondern als Amplifikation des anderen Allels interpretiert wird. Die Amplifikation der Zielsequenzen erfolgte unter Standardbedingungen ausgehend von 4 ng Templat. Zur Bestimmung der Allelverteilung wurden die fluorophormarkierten PCR-Produkte kapillarelektrophoretisch aufgetrennt und analysiert (**2.5.4**). Die Elektropherogramme zeigen die typischen Peakmuster einer heterozygoten Probe mit jeweils zwei Gruppen, wobei jede Gruppe eines der beiden Allele repräsentiert (**Abb. 3-40**). Die unterschiedlich langen PCR-Produkte für die Allele resultieren von *slippage-Ereignissen*, wobei von der Polymerase eine Wiederholungseinheit (hier ein AC-Dinukleotid) zuviel oder zuwenig einbaut wird. Aufgrund der etwas größeren Amplifikationseffizienz für kürzere PCR-Produkte sind in der heterozygoten Normalprobe die Peakflächen für das längere Allel („B") kleiner als für das kürzere Allel („A"), obwohl anzunehmen ist, dass beide Allele in der gleichen Konzentration in der Probe vorliegen (siehe auch **3.4.3.3**).

Die Allelverteilung ergibt sich aus den Verhältnissen der Peakflächen analog zu **Gl. 3.22**, wobei q hier das Verhältnis der Peakflächen von Allel B zu Allel A ist. Für die betrachteten Marker ergeben sich Q-Werte von 3.7 für D17S800 und von 3.2 für D17S1818. Der durch RT-PCR ermittelte Q-Wert für die gleiche Probe beträgt 3.1.

Da beide untersuchten STR-Marker eine sehr ähnliche Ungleichverteilung in der Tumorprobe aufweisen, muss die onkogene Amplifikation mindestens ein Chromosomensegment von 1.4 cM umfassen und nicht ausschließlich den Genlokus betreffen. Die Übereinstimmung der Q-Werte für beide Methoden zeigt weiterhin, dass in der untersuchten Tumorprobe nur eines der beiden Allele von der Amplifikation betroffen war. Diese Aussage ist anhand der Ergebnisse nur einer der Methoden nicht abzuleiten.

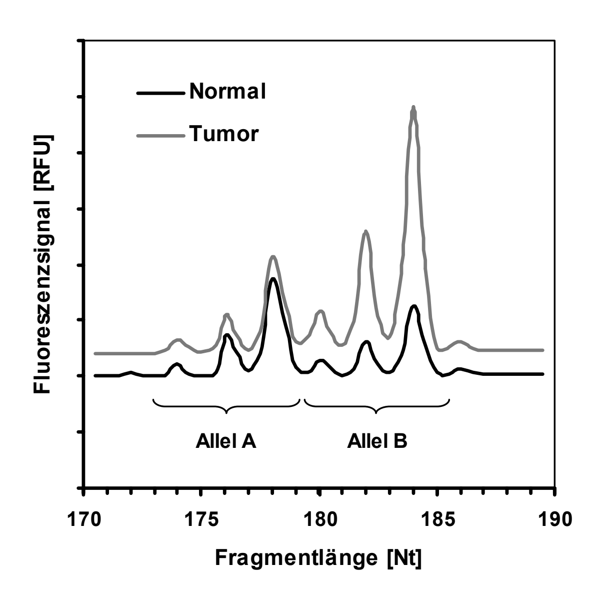


Abb. 3-40 Elektropherogramme von D17S800-PCR-Produkten einer Patientin mit Mammakarzinom. Die den Allelen („A" und „B") zugeordneten Peaks sind gekennzeichnet. In der Tumorprobe (grau) ist die auf Allel A bezogene Menge von Allel B deutlich größer als in der Normalprobe (schwarz).

Die Kombination beider Verfahren zur direkten Bestimmung des Amplifikationsstatus, einerseits durch RT-PCR, andererseits durch Bestimmung der Allelverteilung, kann neue Aufschlüsse über die genetischen Vorgänge während der neoplastischen Transformation von Zellen bei der Entstehung von Tumoren geben. Leider eignen sich jedoch nicht alle Proben zur Untersuchung von Allelverteilungen. So sind z. B. Proben, die für die untersuchten Marker homozygot sind, nicht informativ. Bei großen Ungleichverteilungen kann es zu einem *allelic drop-out* kommen, bei dem das unterpopulierte Allel in der Tumorprobe nicht mehr nachgewiesen werden kann.[131] Weiterhin können theoretisch auch beide Allele von Amplifikationen betroffen sein. Der tatsächliche Amplifikationsstatus ist in diesem Fall größer als das bestimmte Allelungleichgewicht. Ist das Ausmaß der Amplifikationen für beide Allele gleich, findet man sogar kein Allelungleichgewicht, obwohl eine bedeutende Amplifikation vorliegen kann.

Aufgrund dieser möglichen Probleme bei der Untersuchung von Allelverteilungen ist die direkte Quantifizierung der Zielsequenz das geeignetere Verfahren in der medizinischen Diagnostik, wo es lediglich auf die schnelle und sichere Feststellung einer onkogenen Amplifikation ankommt. Das hier entwickelte RT-PCR-basierte Verfahren ist dafür besonders gut geeignet, weil es sehr schnell und einfach durchführbar ist und exakte Ergebnisse liefert.

3.3.3 Bestimmung des Deletionsgrades von Tumorsuppressorgenen

Neben der Amplifikation von Onkogenen ist auch der Verlust funktionsfähiger Tumorsuppressorgene meist an neoplastischen Transformationen beteiligt.[233] Das Tumorsuppressorgen *p53* ist dabei das am besten untersuchte Gen: Eine Inaktivierung von p53 wurde in etwa der Hälfte der untersuchten Fälle nahezu aller Tumortypen diagnostiziert, wobei Deletionen und Punktmutationen die häufigsten Ursachen des Funktionsverlusts sind. In den meisten Tumoren befindet sich ein mutiertes *p53*-Allel, während das andere Allel deletiert ist. Das Mutationsspektrum umfasst Punktmutationen, Insertionen, Deletionen und Genrearrangements.[234] Bei den vorwiegend in Karzinomen auftretenden Punktmutationen von *p53* handelt es sich meist um *missense*-Mutationen, im Gegensatz zu den in anderen Tumorsuppressorgenen wie *RB1* und *APC* hauptsächlich auftretenden *nonsense*-Mutationen.[155-157] Die beschriebenen Mutationen von *p53* in Tumoren sind über das ganze Gen verteilt, häufen sich jedoch in den vier hochkonservierten Regionen II-V innerhalb der Exone 5 bis 8.[234] Zur Diagnose des p53-Funktionsverlusts in Tumorzellen durch den Nachweis von Punktmutationen oder kurzen intragenischen Insertionen bzw. Deletionen muss somit mindestens ein Bereich von 900 Bp in 4 verschiedenen Exonen des Gens sequenziert oder mit einem der anderen in **1.3** angesprochenen Verfahren analysiert werden. Bei der Beschränkung des Mutationsnachweises auf die Kodone der sechs am häufigsten mutierten Aminosäurereste lässt sich nur in etwa 35 % der Fälle ein p53-Funktionsverlust diagnostizieren.[235]

Der Funktionsverlust von p53 in Tumoren lässt sich damit weitaus einfacher und mit größerer Sicherheit mit dem Nachweis einer Alleldeletion feststellen. Dieser kann entweder direkt durch die Quantifizierung der Allelanzahl erfolgen (mittels Karyotypie, FISH oder QPCR) oder indirekt durch die Messung der Allelverteilung polymorpher DNA-Marker (STRs, VNTRs oder RFLPs) innerhalb oder in der Nähe des betreffenden Gens. Beim indirekten Nachweis muss die Person für den untersuchten Marker heterozygot sein. Der Verlust eines Allels zeigt sich in einem Verlust der Heterozygosität für diesen Marker, weshalb diese Untersuchungen auch als *loss of heterozygosity* (LOH)-Analysen bezeichnet werden. Die Nachteile dieses Verfahrens bestehen darin, dass zum einen entsprechende Marker bekannt sein müssen, für welche die untersuchte Person heterozygot ist, und zum anderen, dass die quantitativen Aussagen nicht absolut, sondern nur relativ sind. So kann nicht nur der Verlust des einen Allels, sondern auch die Amplifikation des anderen Allels Ursache einer gefundenen Allelungleichverteilung sein. Die direkten Verfahren erlauben die absolute Bestimmung der Genkopienzahlen je Genom, wodurch Allelverluste zweifelsfrei nachgewiesen werden können, allerdings sind sowohl die Karyotypie als auch die FISH-Techniken vergleichsweise aufwändig in Durchführung und Auswertung, weshalb vorwiegend PCR-basierte Verfahren diagnostische Verwendung finden. Eine Übersicht über ausgewählte Publikationen zum Nachweis von Aneuploidien, Genamplifikationen und Gendeletionen findet sich bei M. HAHN.[205]

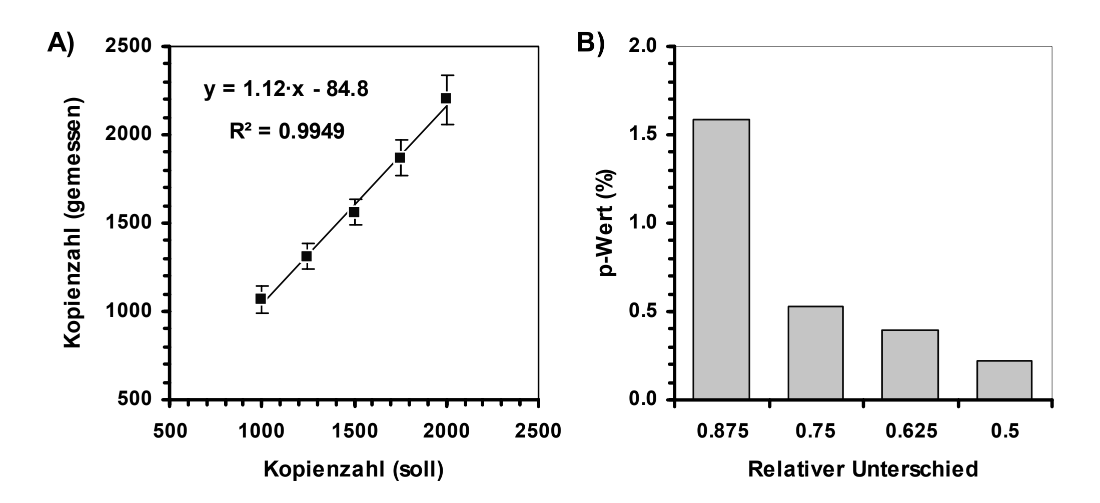


Abb. 3-41 Präzision der Quantifizierung von *p53*. A) Mittelwerte und Standardabweichungen von Fünffachbestimmungen der *p53*-Kopienzahlen. Die Standards enthielten zwischen 8000 und 500 Genomäquivalente je PCR, seriell 1:2 verdünnt. Als Templat diente in allen Fällen Blutlymphozyten-DNA, die Detektion erfolgte mit Hybridisierungssonden. Ausgewertet wurden die Signalverhältnisse von Akzeptor- zu Donorfluorophor. Gleichung der Standardgeraden: $y = -3.45 \cdot x + 35.9$, $R^2 = 0.9944$. Die aus der Steigung der Standardgeraden berechnete Amplifikationseffizienz beträgt 1.95. **B)** Ergebnisse der *pair-wise randomization tests* für alle Kombinationen der Fünffachbestimmungen aus A). Die relativen Unterschiede entsprechen den Verhältnissen der Soll-Kopienzahlen der verglichenen Gruppen. Der p-Wert gibt die Wahrscheinlichkeit an, mit welcher der gemessene Unterschied zufällig ist (Irrtumswahrscheinlichkeit).

Im Folgenden wird ein RT-PCR basiertes Verfahren zur Bestimmung von *p53*-Gendeletionen vorgestellt, welches die Quantifizierung der absoluten Menge von *p53*-Genkopien pro Genom erlaubt.

Das Prinzip ist identisch mit dem in **3.3.1** beschriebenen und beruht auf der Bestimmung eines Q-Wertes. Auch hier dient *IGF-1* als Referenzgen und genomische DNA zur Standardisierung. Im Gegensatz zur Quantifizierung von Genamplifikationen, wo nur selten Konzentrationsunterschiede von weniger als dem Zweifachen aufgelöst werden müssen, ist ein zweifacher Konzentrationsunterschied bei der Quantifizierung von Gendeletionen praktisch die Obergrenze, da zum einen die Wahrscheinlichkeit der Deletion beider Allele in einer Zelle sehr gering ist und zum anderen nicht notwendigerweise alle Zellen in der untersuchten Probe eine Gendeletion aufweisen müssen.[152] Dies stellt höchste Anforderungen an die Präzision des Verfahrens.

Die ermittelten Präzisionen für die Quantifizierung der Genkopienzahlen, ausgehend von 1000 bis 2000 Templatkopien, sind für beide Zielsequenzen vergleichbar. Die Variationskoeffizienten von Fünffachbestimmungen liegen für *IGF-1* im Mittel bei etwa 8.8 %, für *p53* bei circa 5.8 %. Die Signifikanzniveaus für die Unterscheidbarkeit der eingesetzten Templatmengen wurde durch *pair-wise randomization tests*[209] ermittelt (**Abb. 3-41**). Die Wahrscheinlichkeit (p-Wert des Tests), dass der gemessene Unterschied nur zufällig ist, beträgt für 0.875-fache Mengenunterschiede nur etwa 1.6 %, d. h., dass Proben, deren Templatmengen sich lediglich um 12.5 % unterscheiden, noch mit mehr als 98%iger Sicherheit unterschieden werden können.

Da die Fehler von vier unabhängigen Quantifizierungen (des Test- und des Referenzgens jeweils in der Tumor- und in der Normalprobe) in die Berechnung der Q-Werte (**Gl. 3.22**, Seite 109) eingehen, kann diese Präzision für die Q-Werte nicht ganz erreicht werden. Die in Simulationsexperimenten bestimmten Variationskoeffizienten von Q betragen für alle getesteten Deletionswerte (*p53*:*IGF-1* 1, 0.875, 0.75, 0.62 und 0.5) etwa 10 % (**Abb. 3-42**). Ein *pair-wise randomization test*[209] ergibt für die kleinsten getesteten Mengenunterschiede (0.875 bzw. 12.5 %) einen p-Wert von 0.0857 (n=5), für den nächst größeren (0.75) 0.0286. Das bedeutet, dass mit diesem Verfahren bereits der Verlust je eines *p53*-Allels in nur jeder vierten Zelle (25 %) der Tumorprobe mit mehr als 90%iger Sicherheit messbar ist. Hat jede zweite Zelle in der Tumorprobe nur noch ein *p53*-Allel, beträgt die Sicherheit sogar schon mehr als 97 %.

Die Prüfung der klinischen Anwendbarkeit erfolgte an zwei hinsichtlich ihres *p53*-Status bereits mittels konventioneller LOH-Analytik charakterisierten Blasentumorproben[236] mit korrespondierenden Normalgewebeproben, die freundlicherweise von M. HAHN zur Verfügung gestellt wurden.

In beiden Fällen wurden die durch LOH-Analysen ermittelten Q-Werte[236] bestätigt, wobei allerdings die Q-Werte des hier vorgestellten Verfahrens für die Proben etwas größer sind, was ein geringeres Ausmaß von Genverlusten impliziert ($Q_{LOH} : Q_{real\text{-}time}$ 0.61 : 0.75 und 0.68 : 0.80). Bei den veröffentlichten Werten handelt es sich nicht um Q-Werte, welche die Genkopienzahl pro haploidem Genom angeben, sondern um q^{LOH}-Werte, welche ein Maß

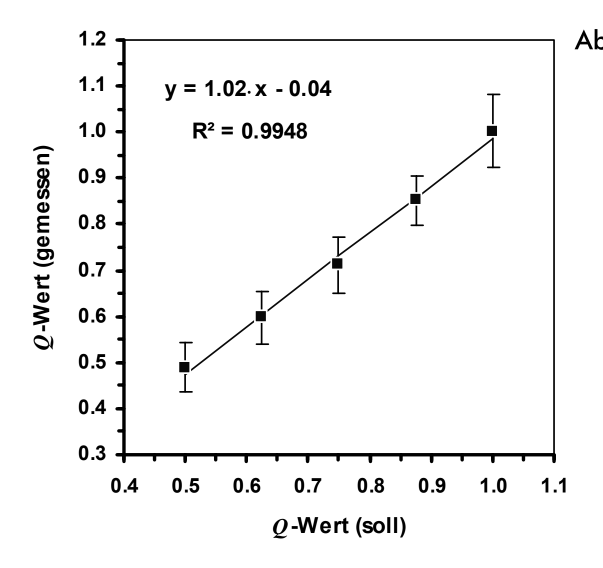


Abb. 3-42 Ergebnis des Simulationsexperiments zum *p53*-Genverlust. Die Q-Werte wurden berechnet nach Gl. 3.22 (Seite 109), wobei hier *p53* das Testgen ist. Die Datengewinnung und Auswertung erfolgten analog zur Bestimmung der simulierten Genamplifikation in 0. Die Templatmenge in der „Normal"-Probe sowie der „Tumor"-Probe beträgt 2000 Genomäquivalente für die *IGF-1*-Messung, für die *p53*-Messung befinden sich 2000, 1750, 1500, 1250 und 1000 Genomäquivalente in den „Tumor"-Proben.

für die Heterozygosität darstellen: Eine vollständig homozygote (bzw. hemizygote) Probe (jede Zelle besitzt nur ein Allel) hat einen q^{LOH} von Null, eine heterozygote Probe mit einer Gleichverteilung beider Allele (jede Zelle besitzt zwei Allele) hat einen q^{LOH} von Eins. Die entsprechenden Q-Werte können daraus abgeleitet werden durch $Q = 0.5 \cdot (q^{LOH}+1)$.

Die hier vorgestellte Methode erlaubt also die einfache und schnelle Quantifizierung von Genverlusten anhand von wenig Probenmaterial auch mit fragmentierter DNA. Da diese Technik von der Mischung der Komponenten bis hin zur Auswertung der Daten voll automatisierbar ist, kann es zukünftig möglich sein, in Hochdurchsatzverfahren bereits während der Tumorprogression diverse Tumortypen auf Veränderungen des Status tumorrelevanter Gene zu untersuchen und so die Basis für genaue molekulare Charakterisierungen und eine exakte Klassifizierung des Tumors zu entwickeln. Einfache Messverfahren, die mit wenig Probenmaterial auskommen, können weiterhin helfen, die mikroevolutiven Prozesse während der Tumorentstehung besser verstehen zu lernen.

3.4 Methodische Entwicklungen

Schlussgelöbnis über Wissenschaft. Wenn ihr es könnt, werdet ihr müssen.

Friedrich Nietzsche

Nachfolgend werden drei neue, RT-PCR-basierte Methoden vorgestellt, die im Rahmen dieser Arbeit entwickelt wurden. Die kompetitive RT-PCR ermöglicht die Quantifizierung durch Einsatz interner Standards (Kompetitoren) mit sehr hoher Präzision, aber eingeschränktem dynamischen Bereich. Mit der allelspezifischen RT-PCR lassen sich die Konzentrationen sehr ähnlicher Sequenzvarianten in Mischproben innerhalb eines sehr großen Konzentrationsbereichs bestimmen. Schließlich wird noch ein RT-PCR-basiertes Verfahren vorgestellt, welches die einfache, schnelle und sehr präzise Bestimmung von Genomgrößen erlaubt.

3.4.1 Verfahren zur Bestimmung von Genomgrößen

Als eine weitere methodische Entwicklung, die gleichzeitig als Anwendungsbeispiel quantitativer RT-PCR dient und damit zu **3.3** überleitet, wird ein Verfahren zur Bestimmung von Genomgrößen dargestellt, mit welchem hier exemplarisch die Größen der Genome dreier eukaryoter Organismen ermittelt wurden: von der einzelligen Bäckerhefe *Saccharomyces cerevisiae* als vollständig sequenzierten Referenzorganismus, vom Platyfisch *Xiphophorus maculatus*, einem wichtigen Tiermodell in der Krebsforschung sowie vom menschlichen Genom, dessen Genomgröße bereits mit vielen unterschiedlichen Verfahren bestimmt wurde und für dessen Wert durch die Ergebnisse der Sequenzierprojekte inzwischen recht verlässliche Schätzungen vorliegen.

In der Literatur ist eine Vielzahl unterschiedlicher Verfahren zur Genomgrößenbestimmung beschrieben, von denen hier nur die prominentesten kurz vorgestellt werden. In ersten biochemischen Arbeiten erlaubten quantitative chemische Analysen, z. B. des nukleären Phosphatgehalts, Rückschlüsse auf den Gehalt an Nukleinsäuren je Zellkern. Die so ermittelten Werte sind jedoch meist zu groß,[237] wahrscheinlich bedingt durch die Anwesenheit nukleärer Ribonukleinsäuren. Als klassisches biophysikalisches Verfahren wird schon seit den späten 1960er Jahren bis auf den heutigen Tag die Analyse von DNA-Reassoziationskinetiken („c_0t-Assay“) mit Hydroxylapatit-Chomatographie zur Trennung einzel- und doppelsträngiger DNA eingesetzt.[238,239] Neuere Techniken basieren auf der Messung der Fluoreszenzintensität DNA-bindender Farbstoffe in durchflußzytometrischen[240,241] oder absorptionszytophotometrischen[242] Analysen nach FEULGEN-Färbung.[243]

Die Kenntnis von Genomgrößen ist in vielen Bereichen der Biologie von Bedeutung. Die Vergleiche der Genomgrößen verschiedener Arten können z. B. Aufschluss über Verwandtschaftsverhältnisse bzw. evolutive Prozesse sowie über Differenzierungsprozesse während der Ontogenese und Veränderungen des nukleären DNA-Gehalts während der Tumorgenese geben.[241,244-246]

3.4.1.1 Prinzip

Das hier vorgestellte Verfahren beruht auf der Quantifizierung genetischer Elemente, die in bekannter Zahl im Genom vorliegen, durch quantitative RT-PCR mit externer Standardisierung. Die zur Quantifizierung eingesetzte Probenmenge kann UV-spektrophotometrisch bestimmt werden. Das Schema des experimentellen Ablaufs ist in **Abb. 3-43** gezeigt.

Bei der für die Quantifizierung in der RT-PCR benötigten Standard-DNA handelt es sich um ein PCR-Produkt der Zielsequenz. Da die Länge des PCR-Produkts ($n_{Bp,PCR}$) bekannt ist, kann ausgehend von der UV-spektrophotometrischen Bestimmung der Massenkonzentration des PCR-Produkts ($c_{m,PCR}$) die Anzahl der PCR-Produktkopien je Volumeneinheit ($c_{K,PCR}$) berechnet werden nach

$$c_{K,PCR}\left[\frac{Kopien}{Volumen}\right]=\frac{c_{m,PCR}\left[\frac{Masse}{Volumen}\right]}{m_{K,PCR}\left[\frac{Masse}{Kopie}\right]} \qquad \text{Gl. 3.23}$$

wobei $m_{K,PCR}$ die Masse eines PCR-Produktmoleküls bezeichnet. Diese lässt anhand der bekannten Produktlänge in Basenpaaren ($n_{Bp,PCR}$) und der Masse eines Basenpaares ($m_{Bp}=M_{Bp}/N_A$; M_{Bp}: molare Masse eines Basenpaars, ≈ 660 g/mol und N_A: Avogadrosche Zahl ≈ $6.022 \cdot 10^{23}$ mol^{-1}) ermitteln:

$$m_{K,PCR}\left[\frac{Masse}{Kopie}\right] = n_{Bp,PCR} \cdot m_{Bp} = n_{Bp,PCR} \cdot \frac{M_{Bp}\left[\frac{Masse}{Mol}\right]}{N_A\left[\frac{Kopie}{Mol}\right]} \qquad \text{Gl. 3.24}$$

Durch die Kenntnis von $c_{K,PCR}$ der Standards kann mittels der quantitativen RT-PCR die absolute Kopienkonzentration der Zielsequenz in der Proben-DNA bestimmt ($c_{K,DNA}$) werden. Die Größe des haploiden Genoms (1n), Γ, in früheren Arbeiten auch als S-Wert bezeichnet, entspricht dem Quotienten der UV-spektrophotometrisch bestimmten Massenkonzentration ($c_{m,DNA}$) und dem Produkt der in der Quantifizierung bestimmten absoluten Kopienkonzentration ($c_{K,DNA}$) mit der Anzahl der Kopien der Zielsequenz im haploiden Genom (f):

$$\Gamma\left[\frac{Masse}{Kopie}\right] = \frac{c_{m,DNA}\left[\frac{Masse}{Volumen}\right]}{f \cdot c_{k,DNA}\left[\frac{Kopien}{Volumen}\right]} \qquad \text{Gl. 3.25}$$

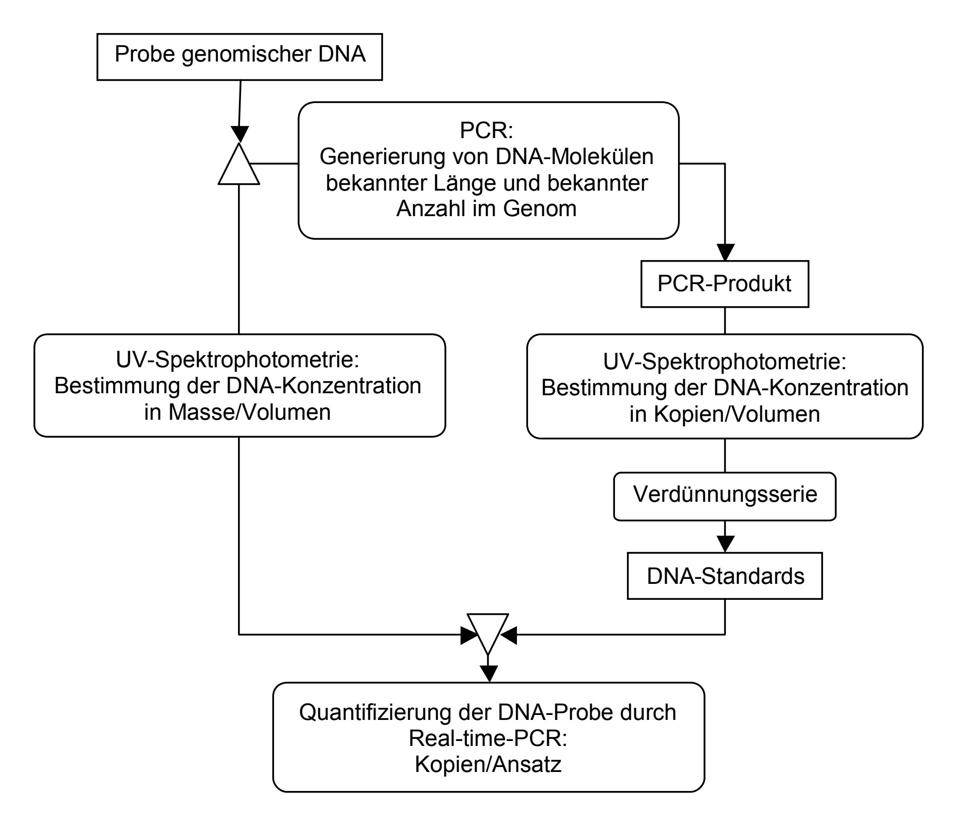


Abb. 3-43 Flussdiagramm zur Genomgrößenbestimmung. Ausgehend von einer Probe genomischer DNA wird zunächst ein PCR-Produkt generiert, welches später als Standard-DNA zur Bestimmung der Kalibrierfunktion eingesetzt wird. Das genaue Verfahren ist im Text beschrieben.

Für den Fall einer Gensequenz, die nur einmal im haploiden Genom vorkommt, wie es für *single-copy*-Gene der Fall ist, gilt $f = 1$, womit sich **Gl. 3.25** vereinfacht zu $\Gamma = c_m/c_k$. Die Zahl der Basenpaare pro haploidem Genom kann aus **Gl. 3.25** berechnet werden, indem Γ auf die Masse eines Basenpaares (m_{Bp}) bezogen wird.

Die Amplifikation des Standard-PCR-Produktes kann mit einem Primerpaar erfolgen, welches die Bindungsstellen für die zur RT-PCR verwendeten Primer einschließt. Bei dieser „*nested*-PCR" liegen dann die Bindungsstellen der zur RT-PCR verwendeten Primer nicht genau am Rand des zu kopierenden Templats, wie es der Fall wäre, wenn zur Generierung der Standards die gleichen Primer verwendet würden. In den ersten Zyklen der Amplifikation kann die Effizienz von so genannten Randeffekten des Templats beeinflusst sein (z. B. durch Ausbildung stabiler Sekundärstrukturen, welche die Hybridisierung der Primer stören). Durch die Verwendung eines Templats mit entsprechenden Randbereichen als Standard-DNA versucht man zu gewährleisten, dass die Amplifikationseffizienz bei Standards (als PCR-Produkt) und zu quantifizierender Proben (als genomischer DNA) möglichst identisch ist, was eine essentielle Voraussetzung für eine exakte absolute Quantifizierung mittels der RT-PCR darstellt.

3.4.1.2 Beispiel *Saccharomyces cerevisiae*

Da das Genom der Bäckerhefe sequenziert ist, liegt für diesen Organismus ein sehr exakter Wert für die Genomgröße vor, weshalb dieser hier als Referenzmodell für das Verfahren zur Genomgrößenbestimmung dienen soll. Für die Untersuchung wurde der Stamm #368 FY 1679 verwendet, der die Grundlage für die Sequenzierungsprojekte war.[247]

Die UV-spektrophotometrisch bestimmte Konzentration der DNA-Probe, isoliert aus den Zellen einer Hefekultur, betrug $c_{m,DNA} = 17.9$ ng/µl. Quantifiziert wurde das *single-copy*-Gen *rps3* für das ribosomale Protein S3. Das nach dem Standard-PCR-Protokoll hergestellte und aufgereinigte PCR-Produkt der Primer RPS-H1/R1 mit einer Länge von 382 Bp wurde UV-spektrophotometrisch bestimmt und mit Reinstwasser auf eine Konzentration von 10^8 K/µl eingestellt.

Zur Quantifizierung der Kopienzahl der Zielsequenz im Hefegenom wurde eine serielle 1:10-Verdünnungsreihe der Standard-DNA (von 10^8 bis 10^2 K/µl) sowie ein Fünffach-Ansatz der genomischen DNA mit dem Primerpaar RPS-H2/R2 im LightCycler amplifiziert. Das 172 Bp lange Produkt wurde mittels SYBR-Green I detektiert. Anhand der Kalibriergeraden ($y = -4.0 \cdot x + 38.7$, $R^2 = 0.9993$) wurde für die Probe eine Konzentration von $c_{K,DNA} = 1.35 \cdot 10^6$ K/µl ± 6 % (MW ± SD) ermittelt. Nach **Gl. 3.25** folgt damit für die (haploide) Genomgröße ein Wert von $\Gamma = 13.26$ fg, was einer Länge von 12.1 MBp entspricht. Dieser Wert stimmt sehr genau mit der durch Sequenzierung bestimmten Genomgröße von 12.069 MBp überein.[247]

3.4.1.3 Beispiel *Homo sapiens*

Die Größe des menschlichen Genoms wurde anhand dreier DNA-Proben aus Blutlymphozyten mit einer Konzentration von jeweils $c_{m,DNA} = 15$ ng/µl bestimmt. Als Standard-DNA

diente das mittels konventioneller PCR nach dem Standard-PCR-Protokoll hergestellte und aufgereinigte PCR-Produkt der Primer QPCR-2a/2b mit einer Länge von 342 Bp.

Gemessen wurde die Kopienzahl des *p53*-Gens in Fünffach-Ansätzen nach Amplifikation mit dem Primerpaar *p53*-H/R und Detektion mit den *p53*-spezifischen Hybridisierungssonden im LightCycler. Ausgewertet wurde das Signalverhältnis von Kanal 2 / Kanal 1. Die ebenfalls in Fünffach-Ansätzen gemessenen Standards enthielten 10^5, 10^4 und 10^3 Kopien des nach dem Standard-PCR-Protokoll hergestellten und aufgereinigten, 125 Bp langen PCR-Produkts der Primer *p53*-H/R.

Anhand der Kalibriergeraden ($y = -3.71 \cdot x + 38.3$, $R^2 = 0.9978$) wurden für die Proben eine mittlere Konzentration von $c_{K,DNA} = 4.69 \cdot 10^3$ K/µl ermittelt. Die Standardabweichungen der 5fach-Bestimmungen liegen zwischen 1 % und 5 %. Die nach **Gl. 3.25** berechnete (haploide) Genomgröße ist $\Gamma = 3.2$ pg bzw. 2.9 GBp. Der Variationskoeffizient der unabhängigen Bestimmungen der drei Proben beträgt 13 %.

Die Angaben in den üblichen Lehrbüchern über die bereits mit vielen unterschiedlichen Verfahren bestimmte Genomgröße des Menschen schwanken zwischen 2.8 und 3.2 GBp.[248-250] Die Schätzungen der Genomgröße, ausgehend von den Sequenzierungsergebnissen des Humanen Genomprojekts, legen eine etwas kleinere Genomgröße nahe, wobei Werte zwischen 2.8 und 3.0 GBp angegeben werden.[250] Der hier bestimmte Wert von 2.9 GBp liegt genau in diesem Bereich und entspricht dem von Celera Genomics angegebenen Wert von 2.9 GBp für die Konsensussequenz des euchromatischen Anteils des Genoms.[58]

3.4.1.4 Beispiel *Xiphophorus maculatus*

Wie noch in **3.3.1** beschrieben wird, ist der ursprünglich in zentralamerikanischen Gewässern beheimatete Platyfisch (*Xiphophorus maculatus*) nicht nur ein beliebter Aquarienfisch, sondern auch ein wichtiges Tiermodell in der Verhaltensforschung und der Krebsforschung.[224-226] Nicht zuletzt aus diesem Grund wurde der nukleäre DNA-Gehalt dieses Fisches schon mehrfach mit unterschiedlichen Methoden bestimmt, deren Ergebnisse hier zum Vergleich herangezogen werden können.

Analog zum Vorgehen für die Hefe wurde auch hier, ausgehend von einer DNA-Probe mit einer Konzentration von $c_{m,DNA} = 16.8$ ng/µl, mit Hilfe der Primer Xp53-H1/R1 zunächst ein 387 Bp langes PCR-Produkt generiert und auf eine Konzentration von 10^8 K/µl eingestellt. Zur RT-PCR wurden die Primer Xp53-H2/R2 eingesetzt, welche ein 89 Bp langes Produkt liefern. Der anhand der Kalibriergeraden ($y = -3.9 \cdot x + 43.0$, $R^2 = 0.9998$) bestimmte Wert für die Konzentration der Probe ist $c_{K,DNA} = 2.78 \cdot 10^4$ K/µl ± 8 % (MW ± SD; n=5).

Für die Größe des haploiden Genoms ergibt sich damit nach **Gl. 3.25** ein Wert von $\Gamma =$ 604 fg bzw. 551 MBp.

In der Literatur finden sich Werte für die Genomgröße von Xiphophorus maculatus zwischen 410 MBp (450 fg) und 550 MBp (600 fg), abhängig von der jeweiligen Untersu-

chungsmethode. Über Reassoziationskinetiken wurde ein Wert von 530 MBp (577 fg) bestimmt, die absorptionszytophotometrische Analyse nach Feulgen-Färbung lieferte einen geringfügig größeren Wert von 540 MBp (595 fg).[251] Das mit der hier vorgestellten Methode gefundene Ergebnis von 550 MBp (604 fg) stimmt auch hier sehr gut mit den Literaturwerten überein.

Die Ergebnisse für die drei untersuchten Arten zeigen, dass mit Hilfe des hier beschriebenen RT-PCR-basierten Verfahrens sehr verlässliche und genaue Aussagen über die Größe von Genomen möglich sind. Im Vergleich zu den in der Literatur beschriebenen Verfahren[238-243] ist hierbei jedoch der experimentelle und zeitliche Aufwand deutlich geringer. Die Genauigkeit dieser Methode hängt entscheidend ab von der Spezifität der PCR-Reaktionen (es dürfen keine Nebenprodukte gebildet werden) sowie der Genauigkeit der UV-spektrophotometrischen Quantifizierung. Die Bildung nicht gewünschter Nebenprodukte bei der Generierung der Standard-DNA führt zu einer Unterschätzung der Genomgröße, da ein zu großer Wert für die Kopienkonzentration der Standards berechnet wird. Fehler in der UV-spektrophotometrischen Quantifizierung können dabei sowohl zur Unter- als auch zur Überschätzung der Genomgröße führen. Hier sollte statt einer Zweipunktbestimmung (bei 260 und 320 nm) ein vollständiges UV-Absorptionsspektrum (mit einem Wellenlängenbereich von 220-320 nm) gemessen werden, um die Qualität der DNA-Lösung beurteilen zu können und nicht aufgrund von Verunreinigungen mit anderen UV-Licht-absorbierenden Substanzen, insbesondere Proteinen, zu große Werte zu messen. Eine Überbestimmung der Konzentration der genomischen DNA-Probe hat auch eine Überschätzung der Genomgröße zu Folge, während die Überbestimmung der Konzentration der Standard-PCR-Produkte zur Unterschätzung der Genomgröße führt.

3.4.2 Kompetitive Real-time-PCR

RT-PCR basierte Methoden zur Quantifizierung von Nukleinsäuren verwenden üblicherweise externe Standards zur Bestimmung der initialen Kopienzahlen bzw. Konzentrationen der Template. Dieses Verfahren hat einen sehr großen dynamischen Bereich, allerdings eine eingeschränkte Präzision: Die Variationskoeffizienten von Einzelbestimmungen liegen nach Herstellerangaben bei bis zu 50 %, so dass in der Regel Mehrfachbestimmungen mit fünf und mehr Reaktionsansätzen je Probe durchgeführt werden müssen, um hinreichende Genauigkeiten zu erzielen. Ein weiterer kritischer Punkt ist, dass zur Quantifizierung die Amplifikationseffizienzen von Standards und unbekannten Proben identisch sein müssen (siehe **0**). Diese Vorausetzung ist bei einer Koamplifikation der Standard- (bzw. Kompetitor-) und Probenemplate im selben Reaktionsansatz bei Verwendung derselben Primer mit maximaler Wahrscheinlichkeit erfüllt, allerdings müssen die PCR-Produkte der Kompetitor- und Probentemplate hierbei analytisch unterscheidbar sein (**1.2.1**). In der RT-PCR kann dazu das Schmelzverhalten entweder der PCR-Produkte mit SYBR-Green I oder der Produkt/Sonden-Hybride analysiert werden.

Die Möglichkeit einer quantitativen Analyse von Schmelzkurven kompetitiv amplifiierter Template unter Verwendung von SYBR-Green I wurde zuerst von RIRIE *et al.*[97] postuliert und

später von AL-ROBAIY *et al.*[217] gezeigt. Hierzu müssen die Unterschiede zwischen Probe und Kompetitor einen messbaren Effekt auf die Schmelztemperatur der jeweiligen PCR-Produkte haben, womit die Wahrscheinlichkeit gering ist, dass die Amplifikationseffizienzen von Probe und Kompetitor hinreichend ähnlich sind.

Um geringere Unterschiede zwischen Kompetitor- und Probentemplat analysieren zu können, bieten sich sequenzspezifische Sonden an, deren Hybride mit den jeweiligen PCR-Produkten unterscheidbare Schmelztemperaturen aufweisen. Ein Prinzip, welches bereits zur Mutationsdetektion und Allelotypisierung eingesetzt wird.[43,252] Die quantitative Analyse erfolgt wie im Fall von SYBR-Green I ebenfalls im Anschluss an die Amplifikation durch eine Schmelzkurvenanalyse.[147,253]

Im Folgenden wird ein Verfahren vorgestellt, welches eine hochgenaue Quantifizierung mit sequenzspezifischen Sonden ermöglicht, wobei die Daten der Amplifikationskurven verwendet werden und somit eine anschließende Schmelzkurvenanalyse nicht mehr notwendig ist. Dabei kommt nur ein Sondenpaar (Donor- und Akzeptorsonde) zum Einsatz, mit dem die PCR-Produkte sowohl der Probe als auch der Standards während der Amplifikation bei unterschiedlichen Messtemperaturen detektiert werden.

3.4.2.1 Prinzip

Das Grundprinzip der quantitativen, kompetitiven PCR ist in **1.2.1** kurz beschrieben. Hierbei wird die Zielsequenz in Kompetition zu einem internen Standard, dem Kompetitor, amplifiziert. Unter der Voraussetzung, dass beide Sequenzen mit derselben Effizienz amplifiziert werden, bleibt das Mengenverhältnis beider Sequenzen während der PCR konstant, wie sich aus **Gl. 1.4** leicht ableiten lässt. Sei N_0 die unbekannte initiale Menge der zu quantifizierenden Zielsequenz und K_0 die bekannte initiale Menge an Kompetitormolekülen, so ist das Verhältnis N_n:K_n beider Mengen nach n Zyklen gleich dem Verhältnis N_0:K_0, womit sich N_0 bestimmen lässt:

$$\frac{N_n}{K_n} = \frac{N_0 \cdot \varepsilon^n}{K_0 \cdot \varepsilon^n} = \frac{N_0}{K_0} \qquad \textbf{Gl. 3.26}$$

Um sicherzustellen, dass die Amplifikationseffizienzen von Zielsequenz und Kompetitor vergleichbar sind, sollte die Kompetitorsequenz der Zielsequenz in Länge und Zusammensetzung möglichst ähnlich sein und mit demselben Primerpaar amplifiziert werden. Idealerweise unterscheiden sich beide Sequenzen nur in einem Basenpaar. Dieser Unterschied macht sie zur Bestimmung des Mengenverhältnisses analytisch unterscheidbar, z. B. durch eine Spaltung mit einer Restriktionsendonuklease, die nur eines der beiden PCR-Produkte spalten kann. Die Spaltprodukte lassen sich aufgrund ihrer geringeren Länge mit gelelektrophoretischen Verfahren einfach vom nicht spaltbaren Substrat trennen. Die Menge beider Produkte kann dann z. B. mit densitometrischen Verfahren bestimmt werden.

Die RT-PCR bietet die Möglichkeit, PCR-Produkte anhand ihrer Amplifikationskurven parallel, sequenzspezifisch und direkt zu quantifizieren, ohne weitere Analyse der Produkte im Anschluss an die PCR. Für die Detektion der beiden, sich idealerweise nur in einer Position

unterscheidenden Sequenzen können zwei unterschiedlich markierte Sonden verwendet werden, wobei die eine zur Probensequenz und die andere zur Kompetitorsequenz komplementär ist und mit dem jeweils anderen PCR-Produkt nur unter Ausbildung mindestens einer Basenfehlpaarung hybridisieren kann. Deren Fluoreszenz kann während der PCR parallel in verschiedenen Kanälen gemessen werden, um je Reaktionsansatz zwei unabhängige Amplifikationskurven zu erhalten, welche die Akkumulation je eines der beiden PCR-Produkte beschreiben (2-Kanal-Messung). Hierbei konkurrieren jedoch die sequenzspezifischen Sonden um die Bindungsstellen auf den PCR-Produkten, die sich ja idealerweise nur minimal unterscheiden, was sehr stringente Hybridisierungsbedingungen erfordert. Um anhand der ΔC_T-Werte von den in den beiden Kanälen aufgezeichneten Amplifikationskurven von Probe und Standard auf das Konzentrationsverhältnis schließen zu können, muss hier die Amplifikationseffizienz bekannt sein. Diese muss durch zusätzliche PCRs, z. B. einer Verdünnungsserie des Standards, hinreichend genau bestimmt werden.

Das Detektionssystem lässt sich vereinfachen, indem statt zweier unterschiedlicher Sonden zur Detektion der beiden Sequenzen nur eine Sonde verwendet wird, die mit einer der beiden amplifizierten Sequenzen eine Basenfehlpaarung ausbildet. Durch diese Basenfehlpaarung erniedrigt sich die Schmelztemperatur des Sonden/Templat-Hybrids um einige Kelvin, was zur Unterscheidung der PCR-Produkte und deren relativer Quantifizierung in einer Schmelzkurvenanalyse genutzt werden kann. Die Amplifikationskurven können hierzu nicht ohne weiteres, wie oben beschrieben, zur Quantifizierung herangezogen werden. Dieses Verfahren der Quantifizierung durch eine Schmelzkurvenanalyse wurde z. B. von LYON *et al.*[147] zur Bestimmung der *HER-2/neu*-Gendosis verwendet, die damit allerdings nur eine Genauigkeit erreichten, die der von Bestimmungen mit externen Standards vergleichbar ist (mit einer Standardabweichung bei der Mehrfachbestimmung von Kopienzahlen von ca. 30 %).

Bei dem hier vorgestellten Verfahren handelt es sich um eine Abwandlung der von LYON *et al.*[147] beschriebenen Methode, wobei einer der Primer hier jedoch im Unterschuss eingesetzt wird. Bei der sogenannten asymmetrischen PCR folgt einer anfänglichen exponentiellen Zunahme an doppelsträngigen PCR-Produkten eine lineare Zunahme an einzelsträngigen PCR-Produkten mit der zu den Sonden komplementären Sequenz. Dadurch wird eine Produktrehybridisierung verhindert und die Signalausbeute gesteigert,[254,255] was die Messgenauigkeit erhöht. Weiterhin wird wegen der Existenz einzelsträngiger Produkte, die selbst nicht mehr amplifiziert werden, die Hybridisierung der Sonden auf diesen Produkten nicht mehr gestört, so dass während der Amplifikation schon je Zyklus zwei Messwerte bei unterschiedlichen Hybridisierungstemperaturen im selben Kanal aufgezeichnet und daraus zwei Amplifikationskurven erstellt werden können (1-Kanal-Messung). Das Signalverhältnis dieser Messwerte entspricht dem Verhältnis von Gesamtproduktmenge zur Menge des Produkts, zu dem die Sonde vollständig komplementär ist (**Abb. 3-44**). Eine Schmelzkurvenanalyse im Anschluss an die PCR ist nicht mehr erforderlich, kann aber durchgeführt werden, um die Konzentrationsverhältnisse anhand der Peakflächenverhältnisse der Peakkurven zu ermitteln.

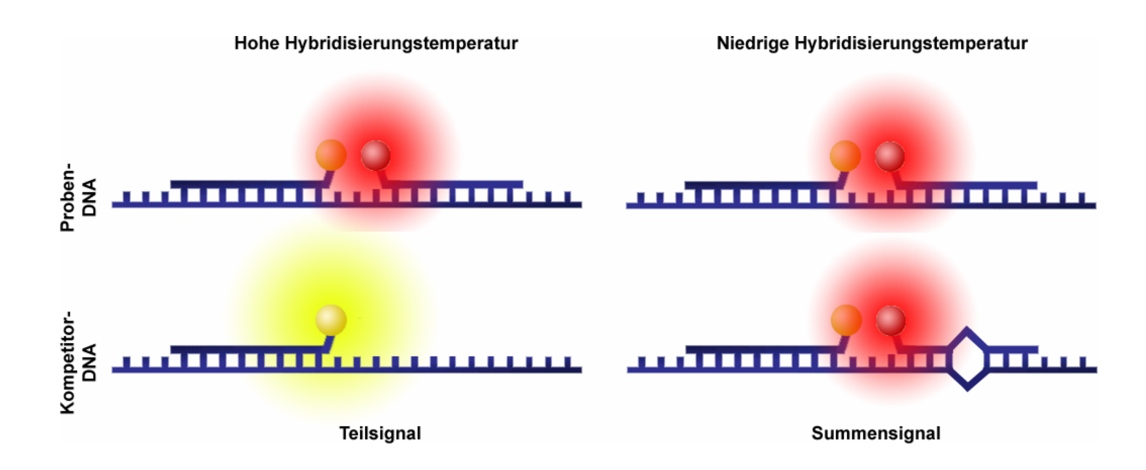


Abb. 3-44 Prinzip der Signalgenerierung bei der kompetitiven RT-PCR mit 1-Kanal-Messung. In dem gezeigten Beispiel ist die Akzeptorsonde komplementär zur Sequenz der Proben-DNA (obere Reihe), während sie bei der Hybridisierung mit der Kompetitor-DNA eine Basenfehlpaarung ausbildet (untere Reihe). Bei hoher, d. h. stringenter Temperatur (linke Seite) bindet die Akzeptorsonde nur am einzelsträngigen PCR-Produkt mit der Probensequenz, das gemessene Fluoreszenzsignal (Teilsignal) ist proportional zur Menge dieses PCR-Produkts. Bei niedriger Temperatur (rechte Seite) hybridisiert die Akzeptorsonde zudem mit dem einzelsträngigen PCR-Produkt des Kompetitors und bei der gemessenen Fluoreszenz handelt es sich um ein Signal, das proportional ist zur Summe beider Produktmengen (Summensignal). Der zum Kompetitorprodukt proportionale Signalanteil ergibt sich aus der Differenz dieser beiden Messwerte.

Da die Fluoreszenzintensitäten für die Teil- und Summensignale bei unterschiedlichen Temperaturen gemessen werden, wird das Teilsignal, welches bei der höheren, stringenten Temperatur gemessen wird, bedingt durch den temperaturabhängigen Quench, im Vergleich zum Summensignal systematisch unterbestimmt. Geht man z. B. von einer äquimolaren Menge beider PCR-Produkte aus, so ist das Summensignal mehr als doppelt so groß wie das Teilsignal. Die korrekten Werte erhält man jedoch einfach durch die Normalisierung der gemessenen Verhältnisse von Teil- zu Summensignal auf das entsprechende Verhältnis der Messwerte von einer bekannten, äquimolaren Mischung beider Produkte.

3.4.2.2 Primer, Sonden und Template

Die Leistungsfähigkeit dieses Verfahrens wurde beispielhaft an der Sequenz des *BCR/ABL*-Fusionstranskripts nachgewiesen. Die Plasmide mit der Wildtypsequenz (im folgenden als Probe bezeichnet) sowie mit einer durch gezielte Mutagenese in zwei Positionen veränderten Sequenz (im folgenden als Kompetitor bezeichnet) wie auch die Primer und Hybridisierungssonden wurden freundlicherweise von A. FREIST zur Verfügung gestellt und sind in ihrer Dissertation beschrieben.[204] Die Positionen der Basenaustausche sowie der Primer und Sonden sind in **Abb. 3-45** dargestellt.

...**ACGTTCCTGATCTCCTCTGA**CTATGAGCGT**GCAGAGTGGAGGGAGAACATCCG**
GGAGCAGCAGAAG**AAGTGTTTCAGAAGCTTCTCCCTGAC**ATCCGTGGAGCTGCA
GATGCTGACCAACTCGTGTGTGAAACTCCAGACTGTCCACAGCATTCCGCTGAC
CATCAATAAGGAAGATGATGAGTCTCCGGGGCTCTATGGGTTTCTGAATGTCAT
CGTCCACTCAGCCACTGGATTTAAGCAGAGTTCA**TTCAGCGGCCAGTAGCATCT**...

Abb. 3-45 Sequenz des amplifizierten Abschnitts des *BCR/ABL*-Fusionstranskripts sowie die Positionen der Primer und Hybridisierungssonden. Grün: Primer bcr(1), hellblau: Primer abl(2), dunkelblau: Donorsonde LC(1), rot: Akzeptorsonde LC(2). Unterstrichen: Position mit Basenaustauschen zwischen Wildtyp- und Kompetitorsequenz. Gezeigt ist die Wildtypsequenz. Bei der Kompetitorsequenz ist das Adenin (A) gegen ein Guanin (G) und das Thymin (T) gegen eine Cytidin (C) ausgetauscht.

Das PCR-Produkt hat eine Länge von 269 Bp. Die Akzeptorsonde ist komplementär zur Probensequenz. Mit der Kompetitorsequenz hybridisiert sie unter Ausbildung zweier zentral gelegener Basenfehlpaarungen (A:C und T:G).

3.4.2.3 PCR-Protokoll mit asymmetrischer Amplifikation und zwei Messpunkten

Die PCR-Ansätze enthalten, abweichend von dem in **2.6.5** beschriebenen Standardansatz für Hybridisierungssonden, den Primer bcr(1) nur in einer Konzentration von 0.05 μM, was einem Zehntel der Konzentration von 0.5 μM des Primers abl(2) entspricht. Donor- sowie Akzeptorsonden werden in einer Konzentration von jeweils 0.2 μM eingesetzt. Als Proben- und Kompetitortemplat, die in unterschiedlichen Mischungsverhältnissen zu den Reaktionsansätzen gegeben werden, dienen die Plasmide mit den inserierten *BCR/ABL*-Sequenzen. Unter den beschriebenen Reaktionsbedingungen beträgt die Schmelztemperatur der Akzeptorsonde von der Probensequenz 60 °C und von der Kompetitorsequenz 70 °C.

Tab. 3-3 Protokoll der kompetitiven RT-PCR mit 1-Kanal-Messung. Wenn nicht anders angegeben, betragen die Temperaturgradienten 20 K/s.

Programm	Wdh.	Segment Nr.	Temp.1 [°C]	Temp.2 [°C]	Step [K]	Dauer [s]	Messpunkt
1 Initiale Denaturierung	1 x	1	95			30	
2 *Touchdown*	10 x	1	95			0	
		2	60	50	1	5	
		3	72			10	
3 Amplifikation	40 x	1	95			0	
		2	50			5	*single*
		3	65			5	*single*
		4	72			10	
4 Schmelzkurve	1 x	1	72			10	*continous*
		2	95*			0	

* Temperaturgradient: 0.2 K/s

Tab. 3-3 zeigt das entsprechend modifizierte PCR-Protokoll. Im Gegensatz zum Standardprotokoll wird in den ersten zehn Zyklen (Programm „*Touchdown*“) die Hybridisierungstemperatur nach jedem Zyklus um 1 K von 60 °C auf 50 °C erniedrigt. In den folgenden 40 Zyklen (Programm „Amplifikation“) wird nach dem Segment zur Hybridisierung mit der Zieltemperatur von 50 °C und einem Messpunkt ein weiteres Segment eingefügt mit einer Zieltemperatur von 65 °C und einem weiteren Messpunkt, so dass dieses Programm zwei Datensegmente enthält, wobei im ersten das Summensignal (Probe + Kompetitor) und im zweiten das Teilsignal der Probe gemessen wird.

Die Auswertung der während der Amplifikation gewonnenen Daten ist ausschließlich mit dem in dieser Arbeit entwickelten Softwareprogramm möglich, da Programme mit mehr als einem Datensegment von der LightCycler-Software nicht korrekt interpretiert werden.

3.4.2.4 Auswertung der Amplifikationskurven

Die Amplifikationskurven aus den Daten der Signalverhältnisse von Kanal 2 / Kanal 1 sowie die sich daraus ergebenden Verhältnisse von Kompetitor zu Probe sind in **Abb. 3-46** gezeigt. Die Menge des Probentemplats ist in jedem Ansatz konstant $2 \cdot 10^6$ Kopien. Gemessen wird jeweils ein Signal bei 50 °C, welches der Summe der Produktmengen von Kompetitor und Probe entspricht (**Abb. 3-46 A**) sowie ein weiteres Signal bei 65 °C, welches proportional zur Produktmenge der Probe ist (**Abb. 3-46 B**). Ein zur Produktmenge des Kompetitors proportionales Signal erhält man durch die Bildung der Differenz dieser beiden Messwerte. Bezogen auf den Signalwert im 40. Zyklus eines Reaktionsansatzes mit vorausgesetzt bekannter Menge Probe (in diesem Beispiel wurde der Wert mit $2.5 \cdot 10^6$ Kopien Kompetitor verwendet, das entspricht der 1.25-fachen Menge der Probe) ergeben sich die in **Abb. 3-46 C** gezeigten normierten Verhältniswerte von Kompetitor zu Probe.

Die Kurven streuen während der exponentiellen Amplifikation bis etwa zum 10. Zyklus stark, konvergieren in den folgenden Zyklen während der linearen Produktzunahme gegen die tatsächlich eingesetzten (eigentlich unbekannten) Verhältniswerte von Kompetitor zu Probe, womit die Menge der Probe bestimmt werden kann. Die Mittelwerte der letzen fünf Zyklen sind (ist / soll): 5.04 / 5; 2.48 / 2.5; 1.27 / 1.25; 0.630 / 0.625 und 0.312 / 0.312. Obwohl es sich um Einfachbestimmungen handelt, weichen die Ergebnisse nur um weniger als 2 % vom richtigen Wert ab.

Zur Berechnung der Verhältniskurven: Sei $F^U{}_{S,i}$ das im i. Zyklus bei 50 °C gemessene Signal (Summensignal) eines Ansatzes U mit unbekannter Probenmenge und $F^U{}_{P,i}$ das im selben Zyklus bei 65 °C gemessene Signal (Probensignal) desselben Ansatzes. Das Kompetitorsignal ergibt sich dann aus der Differenz $F^U{}_{K,i} = F^U{}_{S,i} - F^U{}_{P,i}$ und der unkorrigierte Verhältniswert durch den Quotienten aus Kompetitorsignal und Probensignal: $Q^U{}_i = F^U{}_{K,i} \cdot (F^U{}_{P,i})^{-1}$. Der korrigierte Verhältniswert ergibt sich durch die Normierung von $Q^U{}_i$ auf $Q^R{}_{max} \cdot q^{-1}$ eines Reaktionsansatzes R mit einem bekannten Mengenverhältnis q (wobei $Q^R{}_{max} = Q^R{}_i = F^R{}_{K,i} \cdot (F^R{}_{P,i})^{-1}$ der Referenz im letzen Zyklus):

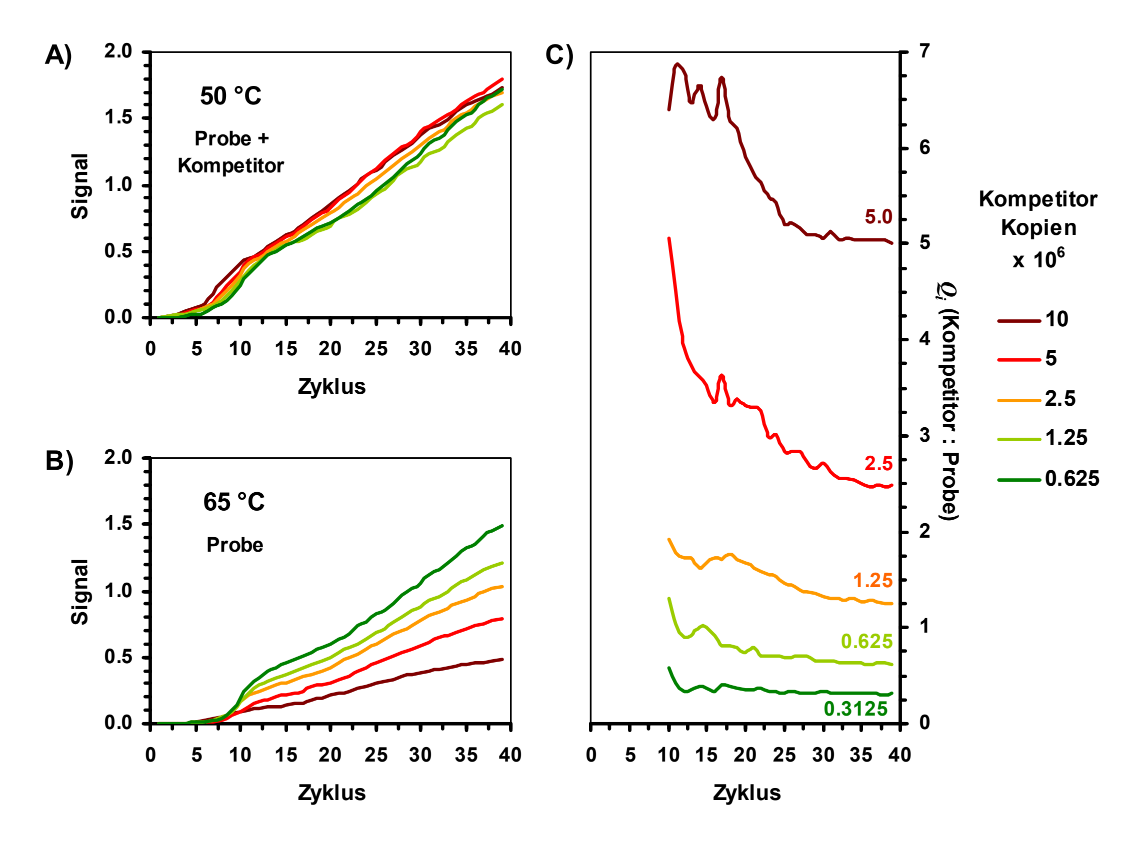


Abb. 3-46 **Amplifikations- und Verhältniskurven einer** kompetitiven RT-PCR. Die Durchführung des Experiments sowie die Berechnung der Verhältniskurven sind im Text beschrieben. Jeder Reaktionsansatz enthält $2{\cdot}10^6$ Kopien des Probentemplats plus die angegebene Menge des Kompetitors. **A)** Amplifikationskurven des Segments Nr. 2 von „Amplifikation". **B)** Amplifikationskurven des Segments Nr. 3 von „Amplifikation". **C)** Verhältniskurven zur Bestimmung der Probenkonzentration. Die Zahlenwerte über den Kurven geben die jeweiligen Sollwerte bzw. die tatsächlichen Mischungsverhältnisse von Kompetitor zu Probe der jeweiligen Reaktionsansätze an (die Probenkopienzahl bträgt jeweisl $2{\cdot}10^6$).

$$Q_i^{U,korr} = \frac{q \cdot Q_i^U}{Q_{max}^R} = q \cdot \frac{F_{K,i}^U / F_{P,i}^U}{F_{K,i}^R / F_{P,i}^R} = q \cdot \frac{F_{K,i}^U \cdot F_{P,i}^R}{F_{K,i}^R \cdot F_{P,i}^U} = q \cdot \frac{F_{K,i}^U \cdot \left(F_{S,i}^R - F_{K,i}^R\right)}{F_{K,i}^R \cdot \left(F_{S,i}^U - F_{K,i}^U\right)} \qquad \text{Gl. 3.27}$$

Statt des Wertes $Q^R{}_{max}$ kann auch der über die letzten 5 Zyklen gemittelte Wert zur Normierung verwendet werden.

3.4.2.5 Auswertung der Schmelzkurven

Die im Anschluss aufgezeichneten Schmelzkurven (**Abb. 3-47**) können ebenfalls zur Bestimmung der Mengenverhältnisse verwendet werden. Die Peakflächen entsprechen hier,

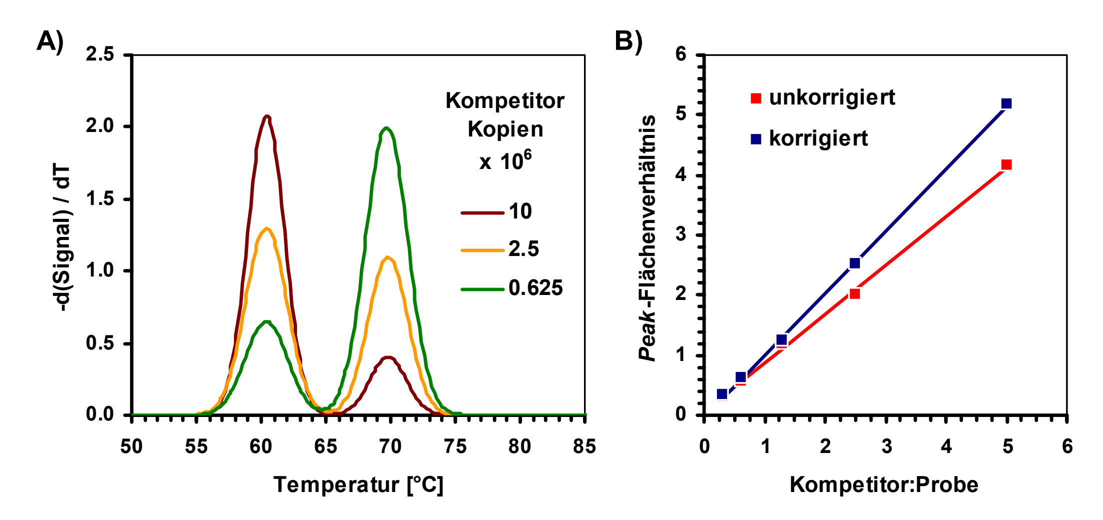


Abb. 3-47 Ergebnis der Schmelzkurvenanalyse nach kompetitiver RT-PCR. Die Durchführung des Experiments ist im Text beschrieben. Ausgewertet wurden die Signalverhältnisse von Kanal 2 und Kanal 1. Jeder Reaktionsansatz enthält $2 \cdot 10^6$ Kopien des Probentemplats plus die angegebene Menge des Kompetitors. **A)** Peakkurven nach Korrektur der temperaturbedingten Quenchs für drei der insgesamt 5 Ansätze. **B)** Verhältnisse der Peakflächen (55-65 °C: Kompetitor; 65-75 °C: Probe) gegen das Verhältnis der eingesetzten Menge von Kompetitor und Probe. Regressionsgeraden: $y = 0.82 \cdot x + 0.06$, $R^2=0.9989$ (unkorrigiert) und $y = 1.04 \cdot x - 0.04$, $R^2=0.9982$ (korrigiert).

nach Anwendung der in **3.2.2.2** beschriebenen Korrektur des temperaturbedingten Quenchs, direkt den Produktmengen von Kompetitor bzw. Probe, so dass zur Auswertung keine Differenzen gebildet werden müssen. Die Normierung auf das Peakflächenverhältnis einer Referenz mit bekannter Probenmenge kann die Richtigkeit der berechneten Ergebnisse erhöhen, sofern die Schmelzkurven nicht zuvor um den temperaturbedingten Quench korrigiert wurden. In diesem Falle werden jedoch aus den in **3.2.2.2** erläuterten Gründen geringe Probenmengen relativ zur Menge Kompetitor unterbestimmt (die Akzeptorsonde hat auf der Probe eine um 10 K höhere Schmelztemperatur als auf dem Kompetitor), wie aus **Abb. 3-47 B** hervorgeht: während die Steigung der Regressionsgeraden für die korrigierten Schmelzkurven einen Wert von 1.04 erreicht, ist ihre Steigung für die unkorrigierten Schmelzkurven mit 0.81 zu gering.

Der mittlere Variationskoeffizient bei der Bestimmung der Verhältnisse von Kompetitor zu Probe anhand der Peakflächenverhältnisse liegt für die unkorrigierten Schmelzkurven bei 11.6 % und für die korrigierten bei 2.9 %. Die über die Schmelzkurvenanalyse und Quenchkorrektur erreichbare Genauigkeit entspricht damit näherungsweise der anhand der Amplifikationskurven erreichbaren Genauigkeit.

3.4.2.6 Demonstration der erreichbaren Genauigkeit

Zur Demonstration der hohen, mit diesem Verfahren erreichbaren Genauigkeit bei der auf den Amplifikationskurven basierenden Quantifizierung (**3.4.2.4**) ist in **Abb. 3-48** das Er-

gebnis einer kompetitiven RT-PCR in einem engen Konzentrationsbereich des Kompetitors um den Äquivalenzpunkt gezeigt.

Der Versuch wurde wie oben beschrieben durchgeführt mit jeweils $2 \cdot 10^6$ Kopien Probentemplat je Reaktion. Die Menge Kompetitor beträgt zwischen $2 \cdot 10^6$ und 10^6 Kopien je Reaktion. Ausgewertet wurden jeweils die mittleren Werte des Signalverhältnisses von Kanal 2 zu Kanal 1 (Ch.2/Ch.1) der Zyklen 35-40 jeder Amplifikationskurve. Als Referenzwert diente Q^R_{max} der Reaktion mit äquimolaren Mengen an Kompetitor und Probe.

Die Standardabweichungen der Dreifachbestimmungen schwanken zwischen 0.013 und 0.022, die Variationskoeffizienten liegen zwischen 0.3 % und 5.7 % mit einem MW von 2.6 %. Die Standardabweichung für den Reaktionsansatz mit äquimolaren Mengen von Kompetitor und Probe ist Null, da dieser zur Normierung verwendet wurde. Entsprechend den Standardabweichungen können mit diesem Ansatz unter idealen Bedingungen noch Probenmengen anhand von Einzelbestimmungen signifikant unterschieden werden, die sich um weniger als 6% unterscheiden. Damit ist gezeigt, dass dieses im Vergleich zur konventionellen kompetitiven quantitativen PCR sehr einfache Verfahren prinzipiell funktioniert und unter idealen Bedingungen sehr genaue Ergebnisse liefert. Der hier bestimmte Wert der Genauigkeit dürfte jedoch bei Verwendung von genomischer DNA als Templat und geringeren Templatmengen nicht ganz erreicht werden.

Die Ergebnisse sind vergleichbar mit den publizierten Daten von LYON *et al.*,[147] die anhand der Schmelzkurvenanalyse sehr ähnlicher Experimente mit humaner genomischer DNA als Probe und einem PCR-Produkt als Kompetitor gewonnen wurden. Dort wurde ein größerer Bereich von Mischungsverhältnissen untersucht, wobei eine lineare Abhängigkeit der Peakflächenverhältnisse nur über den auch hier in **3.4.2.5** untersuchten Bereich von etwa einer Größenordnung gegeben war. In diesem Bereich lagen die von LYON *et al.*[147] ermittelten Variationskoeffizienten der Mittelwerte (n=20) bei ca. 20 %, was vergleichbar ist mit der Genauigkeit, die in Experimenten mit externer Standardisierung erreicht werden kann, aber deutlich unter der Genauigkeit liegt, die mit dem in dieser Arbeit beschriebenen Verfahren der kompetitiven PCR erreicht wurde.

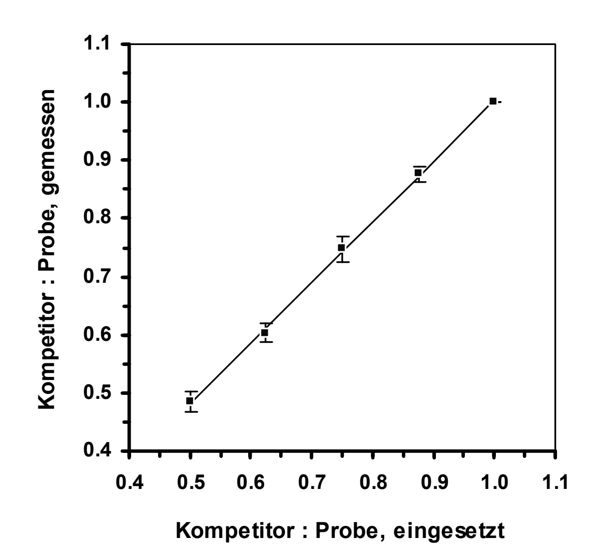


Abb. 3-48 Ergebnis einer kompetitiven RT-PCR. Das Experiment ist im Text beschrieben. Gezeigt sind die Mittelwerte und Standardabweichungen von Dreifachbestimmungen. Gleichung der Regressionsgeraden: $y = 1.04 \cdot x - 0.04$, R^2: 0.9992.

3.4.2.7 Einfluss ungleicher Amplifikationseffizienzen von Proben und Standards auf die Ergebnisse von Quantifizierungen

Die exakte Quantifizierung, sowohl unter Verwendung externer als auch interner Standards (wie bei der hier beschriebenen kompetitiven PCR), erfordert identische Amplifikationseffizienzen für die Proben- und die Standardsequenzen.[17,20,21,256] Dazu dürfen sich die Reaktionsbedingungen, inklusive der Primer-, Sonden- und Templatsequenzen, für die Proben und die Standards idealerweise nicht unterscheiden. Bei der Verwendung externer Standards können proben-, pipettier- und gerätebedingte Variationen wie z. B. das Vorhandensein von Polymeraseinhibitoren, Konzentrationsschwankungen und Temperaturinhomogenitäten die Effizienzen in den unterschiedlichen Reaktionsgefäßen beeinflussen, während bei der Verwendung interner Standards die Sequenzen notwendigerweise nicht identisch sein können und sich so die Amplifikationseffizienzen von Standard und Probe im selben Reaktionsgefäß unterscheiden können.

Eine quantitative Aussage über den relativen Fehler aufgrund nicht-identischer Amplifikationseffizienzen lässt sich aus dem Verhältnis der gemessenen zur tatsächlich eingesetzten initialen Templatmenge ableiten. Dies wird im Folgenden zunächst für die externe, anschließend für die interne Standardisierung hergeleitet.

Bei der Verwendung externer Standards wird die initiale Templatmenge einer Probe P anhand ihres C_T-Werts und der durch die Standards definierten Kalibrierfunktion (**Gl. 1.8**) gewonnen. Bei einer exponentiellen Amplifikation (**Gl. 1.4**) der Probe mit einer Effizienz von ε_P erreicht die Menge ihres PCR-Produkts ausgehend von $N_{0,P}$ initialen Templatkopien den Schwellenwert S gemäß **Gl. 1.6** nach $C_{T,P}$ Zyklen, wobei

$$C_{T,P} = -\frac{1}{\log(\varepsilon_P)} \cdot \log(N_{0,P}) - \frac{\log(A)}{\log(\varepsilon_P)} \qquad \textbf{Gl. 3.28}$$

mit $A = p \cdot S^{-1} = \log(p) - \log(S)$, siehe **Gl. 1.6**.

Die aus $C_{T,P}$ mit der Kalibrierfunktion bestimmte initiale Menge $M_{0,P}$ ergibt sich durch Einsetzen in **Gl. 1.8** nach logarithmieren: $\log(M_{0,P}) = (C_{T,P} + b) \cdot m^{-1}$. Resubstitution der Parameter m und b und Vereinfachung ergibt

$$\log(M_{0,P}) = \frac{\log(\varepsilon)}{\log(\varepsilon_P)} \cdot \left(\log(N_{0,P}) + \log(A)\right) - \log(A) \qquad \textbf{Gl. 3.29}$$

Für den Fall, dass die Effizienzen für Standards und Probe identisch sind ($\varepsilon_P = \varepsilon$) folgt wie erwartet $\log(M_0) = \log(N_0)$. Für $\varepsilon_P \neq \varepsilon$ ergibt sich jedoch ein komplizierter Zusammenhang von M_0 und N_0, der nicht weiter vereinfacht werden kann. Mit nachfolgenden Randbedingungen wird jedoch die Interpretation des Zusammenhangs einfacher: Wie in **Gl. 1.5** beschrieben ist $S = p \cdot N_{C_T}$ bzw. $p \cdot S^{-1} = N_{C_T}^{-1} = A$, also ist $\log(A) = -\log(N_{C_T})$. Damit kann **Gl. 3.29** auch geschrieben werden als

$$\log(M_{0,P}) = \frac{\log(\varepsilon)}{\log(\varepsilon_P)} \cdot \left(\log(N_{0,P}) - \log(N_{C_T})\right) + \log(N_{C_T}) \qquad \textbf{Gl. 3.30}$$

Nun ist N_{C_T} gerade die Menge des PCR-Produkts am Schwellenwert. Der Fehler ist also nicht nur abhängig vom Unterschied der Effizienzen, sondern auch vom Schwellenwert. Je niedriger der Schwellenwert, desto kleiner der durch die Effizienzen verursachte Fehler (für $N_{C_T} = N_{0,P}$ ist $\log(M_{0,P}) = \log(N_{0,P})$). Umgekehrt lässt sich auch argumentieren, dass bei einem gegebenen Schwellenwert der Fehler abhängig ist von $N_{0,P}$ und umso geringer wird, je größer $N_{0,P}$ (bzw. kleiner $C_{T,P}$) wird. Das ist leicht einzusehen, da Abweichungen in den Effizienzen ja in jedem Zyklus den Unterschied zwischen der bei identischen Effizienzen erwarteten Produktmenge zur tatsächlich vorhandenen Produktmenge vergrößern.

Bei der Verwendung interner Standards bzw. Kompetitoren wird die initiale Templatmenge einer Probe P normalerweise anhand des Äquivalenzpunkts bestimmt, an dem die detektierten Produktmengen von Probe und Kompetitor K gleich sind: $N_P = N_K$. Es kann auch einfach das Verhältnis $q = N_P : N_K$ bestimmt werden, um daraus $N_{0,P}$ zu bestimmen ($N_{0,P} = q \cdot N_{0,K}$). Aus **Gl. 1.4** für die Produktmengen nach n Zyklen mit Amplifikationseffizienzen von ε_P für die Probe und ε_K für den Kompetitor folgt dafür

$$q = \frac{N_{0,P} \cdot (\varepsilon_P)^n}{N_{0,K} \cdot (\varepsilon_K)^n} \qquad \text{Gl. 3.31}$$

Der theoretisch richtige Wert von $N_{0,P}$ ist damit $N_{0,P} = q \cdot N_{0,K} \cdot (\varepsilon_K)^n \cdot (\varepsilon_P)^{-n}$. Dieser kann jedoch nicht ohne weiteres so berechnet werden, da die Effizienzen ε_P und ε_K nicht bekannt sind. Wird jedoch angenommen, dass Probe und Kompetitor mit der selben Effizienz $\varepsilon = \varepsilon_P = \varepsilon_K$ amplifiziert werden, vereinfacht sich **Gl. 3.31** zu

$$q = \frac{N_{0,P}}{N_{0,K}} \qquad \text{Gl. 3.32}$$

und entsprechend ergibt sich die initiale Probenmenge zu $N_{0,P} = q \cdot N_{0,K}$.

Der relative Fehler in der Bestimmung von $N_{0,P}$ ergibt sich aus dem Verhältnis dieses Wertes zum theoretisch richtigen Wert und beträgt

$$\frac{q \cdot N_{0,K}}{q \cdot N_{0,K} \cdot (\varepsilon_K)^n \cdot (\varepsilon_P)^{-n}} = \left(\frac{\varepsilon_P}{\varepsilon_K}\right)^n \qquad \text{Gl. 3.33}$$

Auch hier ist der Fehler abhängig von der Zyklenzahl, nach welcher die Produkte schließlich detektiert werden. Der Fehler ist umso geringer, je früher die PCR vor Erreichen der Plateauphase abgebrochen wird. Danach vergrößert sich der Fehler nicht weiter, da dort $\varepsilon_P = \varepsilon_S = 1$.

Die Auswertung anhand der Verhältniskurven bei der kompetitiven RT-PCR (**3.4.2.4**) kann damit aufgrund der geringeren Anzahl betrachteter früherer Zyklen mit exponentieller Amplifikation von Vorteil gegenüber der Auswertung durch Schmelzkurvenanalyse (**3.4.2.5**) sein, wenn nicht sichergestellt ist, dass Standard und Probe mit identischer Effizienz amplifiziert werden.

3.4.3 Allelspezifische Real-time-PCR

Die quantitative Analyse genetischer Chimären oder Mischproben ist von vielfältiger Bedeutung in Biologie und Medizin (siehe **1.4**). DNA-Analyseverfahren beruhen auf der vergleichenden quantitativen Untersuchung individuell unterscheidbarer Sequenzabschnitte wie z. B. bei den polymorphen Segmenten der Immunglobulingene, RFLPs, VNTRs und STRs oder der Quantifizierung einer spezifischen Sequenz, die ausschließlich vom Donor der chimären Probe stammen kann, wie z. B. Y-chromosomenspezifische Sequenzen bei einem Geschlechtschimärismus. Auch die in **3.4** beschriebenen Methoden der kompetitiven RT-PCR können prinzipiell zur Chimärenanalytik bzw. der quantitativen Analyse von Mischproben eingesetzt werden. Allen genannten Verfahren gemeinsam (mit Ausnahme der Y-chromosomenspezifischen PCR)[123] ist der sehr begrenzte dynamische Messbereich der Mischungsverhältnisse von maximal 1:20,[257] was bedeutet, dass ein Allel bzw. eine Sequenzvariante mit einem Anteil von weniger als einem Zwanzigstel der Gesamtmenge nicht mehr nachgewiesen werden kann.

Der dynamische Bereich kann durch die selektive Amplifikation und Quantifizierung der im Unterschuss befindlichen Sequenzvariante, z. B. durch eine allelspezifische PCR, erhöht werden. Die spezifische Amplifikation war eine der ersten beschriebenen Anwendungen der PCR überhaupt.[92,258] Die Diskriminierung der nicht zu amplifizierenden Sequenzvariante wird üblicherweise durch die Spezifität eines der beiden Primer erreicht, dessen 3′-Terminus genau auf der polymorphen Position endet. Ein Primer, der unter Ausbildung einer 3′-terminalen Basenfehlpaarung mit dem Templat hybridisiert, kann von der *Taq* DNA-Polymerase idealerweise nicht verlängert werden. Praktisch ist dieses Ereignis jedoch nicht ganz auszuschließen, so dass auch in Abwesenheit des spezifischen Templats ein PCR-Produkt gebildet wird. Das erreichbare Diskriminierungspotenzial schwankt stark und ist von den involvierten Basen, dem Sequenzkontext und den genauen Versuchsbedingungen abhängig.[259,260]

Im Folgenden wird anhand zweier Insertions-/Deletions-Polymorphismen demonstriert, wie Sequenzvarianten durch die Kombination einer allelspezifischen PCR mit der Real-time-Detektion durch SYBR-Green I einfach identifiziert und mit einem hohen dynamischen Bereich quantifiziert werden können.

3.4.3.1 Prinzip

Bei der allelspezifischen PCR wird die Sequenz eines Primers so gewählt, dass sein 3′-Terminus bei der Hybridisierung auf der polymorphen Stelle des Templats zu liegen kommt. Dieser sogenannte selektive Primer wird von der *Taq* DNA-Polymerase idealerweise nur dann verlängert, wenn er 3′-terminal zum Templat komplementär ist, was nur für eine der im Reaktionsansatz befindlichen Sequenzenvarianten der Fall ist. Die Template mit anderen Sequenzen werden somit nicht kopiert. Da in Folge der Diskriminierung spezifisch amplifizierte Templat ungestört repliziert wird, kann dessen initiale Konzentration nach Detektion mit SYBR-Green I oder anderen Sondensystemen anhand des C_T-Wertes der Amplifikati-

onskurve wie in **0** beschrieben innerhalb eines großen dynamischen Bereiches bestimmt werden.

Tatsächlich jedoch verbleibt immer eine Restwahrscheinlichkeit, dass auch das zu diskriminierende Templat kopiert wird. Die Größe dieser Wahrscheinlichkeit hängt vom Sequenzkontext und den Versuchsbedingungen ab und definiert die Grenze des dynamischen Bereichs quantifizierbarer Mischungsverhältnisse der Sequenzvarianten. Eine Wahrscheinlichkeit von 10^{-3} für die Verlängerung des selektiven Primers auf dem zu diskriminierenden Templat bedeutet dabei, dass in jedem Zyklus durchschnittlich jedes tausendste nichtkomplementäre Templatmolekül kopiert wird. Da hierbei der Primer selbst Teil der Kopie ist, besitzen die in den folgenden Zyklen daran synthetisierten PCR-Produkte die zum selektiven Primer vollständig komplementäre Sequenz und werden genau wie die Produkte des spezifischen Templats amplifiziert. Das Ergebnis einer Quantifizierung ist in diesem Falle nur verwertbar, wenn das Mengenverhältnis zwischen dem nachzuweisenden und dem zu diskriminierenden Templat deutlich größer ist als die Wahrscheinlichkeit der Replikation des „falschen" Templats. Diese Grenze kann bei der RT-PCR leicht anhand der Amplifikationskurve einer Negativkontrolle bestimmt werden. Die Negativkontrolle enthält die gleiche Menge DNA-Kopien wie die Probe, jedoch ausschließlich nicht vollständig zum selektiven Primer komplementäres Templat. Ist der C_T-Wert der Amplifikationskurve einer Probe nur weniger als einen Zyklus kleiner als der C_T-Wert der Negativkontrolle, können keine sicheren qualitativen oder quantitativen Aussagen mehr über das Vorhandensein des „richtigen" Templats getroffen werden.

In den folgenden Abschnitten wird das Verfahren exemplarisch an zwei Insertions-/ Deletions-Polymorphismen demonstriert: einer 10 Bp-Variation im Promotor des Gens für den Koagulationsfaktor VIIc (*F7c*) [261] und den 4G/5G-Polymorphismus im Promotor des Gens für den Plasminogenaktivator-Inhibitor-1 (*PAI-1*) [262]. Beide Polymorphismen haben klinische Relevanz durch eine Assoziation mit Risikofaktoren: Das Vorhandensein der 10Bp-Insertion im *F7c*-Promotor bewirkt eine Reduktion der F7c-Plasmakonzentration [263] und ist assoziiert mit einem erhöhten Myokardinfarktrisiko bei Patienten mit koronarer Arteriosklerose;[264] bei Trägern des 4G-Allels des Polymorphismus im PAI-1-Promotor sind die Plasmakonzentrationen von PAI-1 erhöht, sein Vorhandensein dieses Allels wird diskutiert mit einem erhöhten Risiko für familiär bedingten Herzinfarkt [265,266] und schwerer Präeklampsie.[51] Das hier entwickelte System kann somit direkt Einsatz in der klinischen Diagnostik zur Genotypisierung dieser Polymorphismen finden, wobei es schneller und einfacher durchführbar ist als die bereits in anderen Publikationen beschriebenen Verfahren.[101,267,268]

Die Positionen der verwendeten Primer für die konventionelle PCR sowie die allelspezifische RT-PCR sind in **Abb. 3-49** (*F7c*) und **Abb. 3-50** (*PAI-1*) gezeigt.

Während sich die Allele des *F7c*-Polymorphismus wegen seiner relativ großen Insertion theoretisch einfach durch eine allelspezifische PCR diskriminieren lassen sollten, stellt der *PAI-1*-Polymorphismus eine schwierige Zielsequenz dar, da der Polymorphismus hier nur ein einziges G:C-Basenpaar betrifft, welches zudem noch Teil einer Serie von G:C-Basenpaaren ist, also einem für die Primerbindung problematischen Sequenzkontext.

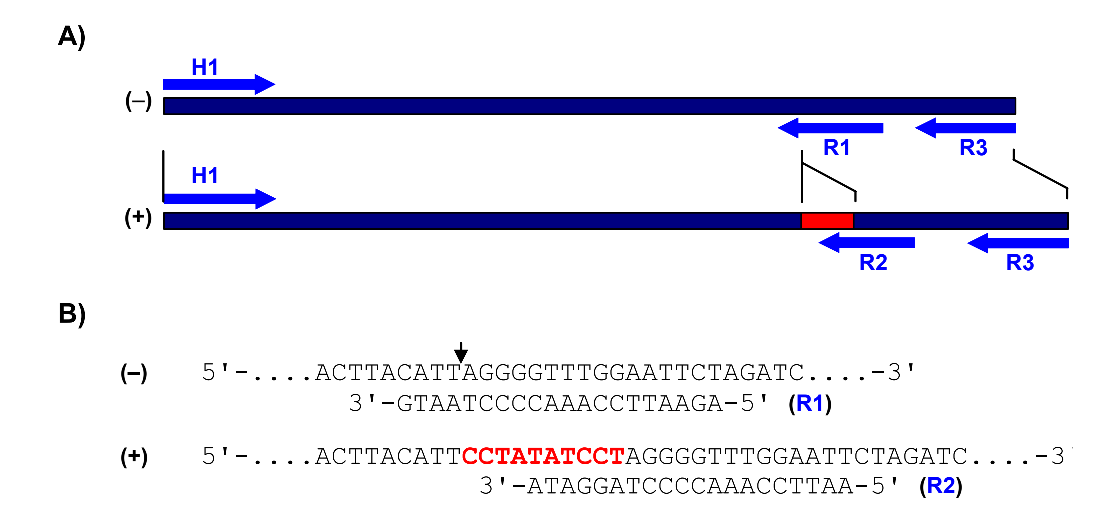


Abb. 3-49 Positionen der Primer **zur Analyse des 10Bp-Polymorphismus im Promotor des Faktors VIIc. A)** (–): Allel ohne Insertion, (+): Allel mit Insertion (rot). Blaue Pfeile: Primer. Die Pfeilspitzen zeigen in 3'-Richtung. Die Längen der PCR-Produkte sind 173 Bp (+) und 163 Bp (–) für das Primerpaar H1/R3, 138 Bp für das (–)-allelspezifische Produkt des Primerpaares H1/R1 und 146 Bp für das (+)-allelspezifische Produkt des Primerpaares H1/R2. **B)** Region der Insertion mit den Bindungsstellen der allelspezifischen Primer. Die Insertionsstelle im (–)-Allel ist durch den Pfeil markiert. Im (+)-Allel ist die Insertion rot hervorgehoben.

Die Sequenzen der allelspezifischen Primer PAI-H1 und PAI-H2 sind nicht vollständig komplementär zur Sequenz ihrer Bindungsstellen auf dem Templat. Beide enthalten an der Position -5 statt eines Guanins ein Adenin, welches bei der Hybridisierung mit dem Cytosin im Templatstrang eine A:C-Basenfehlpaarung ausbildet. Dadurch wird das gesamte 3'-Ende der Primer destabilisiert, was den Effekt einer weiteren 3'-terminalen Basenfehlpaarung mit dem zu diskriminierenden Templatstrang vergrößert. Durch die Einführung dieser obligatorischen Basenfehlpaarung kann die Spezifität bzw. Diskriminierung eines allelspezifischen Primers erhöht werden. Üblicherweise wird dazu die 3'-penultimate oder antepenultimate Position (-2 bzw. -3) für die Basenfehlpaarung gewählt.[269,270] Vergleichend getestete Primer mit obligatorischen Basenfehlpaarungen an den Positionen -2, -3 und -4 zeigten jedoch ein deutlich geringeres Diskriminierungspotenzial als der hier verwendete Primer mit einer Basenfehlpaarung an der Position -5.

Zur Herstellung von Mischproben mit definierten Allelverhältnissen zur weiteren Analyse müssen die Genotypen der vorliegenden Proben bekannt sein. Dazu wurde zunächst eine Genotypisierung des Probenmaterials durchgeführt. Anschließend wurden die dynamischen Bereiche für die Quantifizierung der Allele in Proben mit definierten Mischungsverhältnissen bestimmt, wobei Ergebnisse der allelspezifischen RT-PCR mit denen der kapillarelektrophoretischen Analyse, einer etablierten Referenzmethode zur Genotypisierung und Quantifizierung von Allelungleichverteilungen, verglichen wurden.

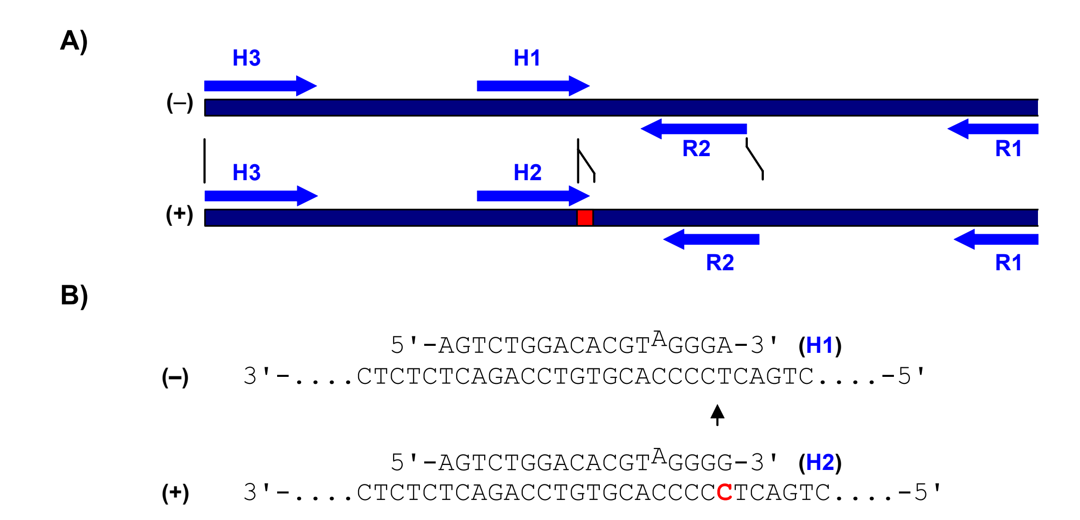


Abb. 3-50 Positionen der Primer **zur Analyse des 4G/5G-Polymorphismus im** *PAI-1*-**Promotor.** A) (–): Allel ohne Insertion, (+): Allel mit Insertion (rot). Blaue Pfeile: Primer. Die Pfeilspitzen zeigen in 3'-Richtung. Die Längen der PCR-Produkte sind 124 Bp (+) und 123 Bp (–) für das Primerpaar H3/R2, 134 Bp für das (–)-allelspezifische Produkt des Primerpaares H1/R2 und 135 Bp für das (+)-allelspezifische Produkt des Primerpaares H2/R2. B) Region der Insertion mit den Bindungsstellen der allelspezifischen Primer. Die Insertionsstelle im (–)-Allel ist durch einen Pfeil markiert. Im (+)-Allel ist die Insertion rot hervorgehoben. Die allelspezifischen Primer besitzen an der Position –5 ein Adenin, welches bei der Hybridisierung mit beiden Allelen eine A:C-Basenfehlpaarung ausbildet.

3.4.3.2 Genotypisierung der Proben

Als Grundlage für die Untersuchungen zur Chimärenanalytik wurden DNA-Proben von 91 gesunden kaukasischen Personen hinsichtlich dieser Polymorphismen charakterisiert, um daraus die zu analysierenden Mischproben mit definierten Mischungsverhältnissen der Allele herstellen zu können. Die Genotypisierung der Proben erfolgte durch Amplifikation der jeweiligen Zielsequenzen mit nicht-diskriminierenden Primerpaaren (siehe **2.6.4**), von denen einer 5'-terminal FAM-markiert war. Die Primer schließen den Polymorphismus ein und liefern in Abhängigkeit des Vorhandenseins der Insertion im Templat unterschiedlich lange, fluorophormarkierte PCR-Produkte. Die Längen der entstandenen PCR-Produkte werden anschließend in der Kapillarelektrophorese analysiert (siehe **2.5.4**).

Zur *F7c*-Genotypisierung wurde das Primerpaar F7c-H1-FAM/R3 verwendet. Von den untersuchten Proben waren 59 homozygot ohne Insertion (–/–), eine homozygot mit Insertion (+/+) und 31 heterozygot (+/–). Als Allelfrequenzen ergaben sich aus diesen Werten 0.18 für das (+)-Allel und 0.82 für das (–)-Allel. Sehr ähnliche Werte von 0.12(+) zu 0.88(–) wurden von HUMPHRIES *et al.* bereits für die britische Population bestimmt.[263]

Der 4G/5G-Polymorphismus in *PAI-1* wurde unter Verwendung des Primerpaares PAI-H3/R2-FAM untersucht. Unter den 59 getesteten Proben befanden sich 24 homozygote

ohne Insertion (–/–), 30 homozygote mit Insertion (+/+) und 37 heterozygote (+/–). Als Allelfrequenzen ergaben sich entsprechend 0.53 für das (+)-Allel und 0.47 für das (–)-Allel, was mit den von YE *et al.* bestimmten Frequenzen (0.46 (+) zu 0.54 (–)).[266] für die britische Population gut übereinstimmt.

3.4.3.3 Kapillarelektrophoretische Analyse von Mischproben

Die Mischproben wurden durch Zusammenfügen für den jeweiligen Polymorphismus homozygoter Proben hergestellt. Diese Proben wurden nach PCR-Amplifikation in der Kapillarelektrophorese analysiert (**Abb. 3-51**).

Die Peakflächen in den Elektropherogrammen entsprechen den PCR-Produktmengen. Unter der Annahme identischer Amplifikationseffizienzen für beide Allele entspricht das Verhältnis der Flächen der beiden Peaks dem Verhältnis der initialen Templatmengen. Kürzere Allele (–) werden jedoch in der Regel mit etwas höherer Effizienz amplifiziert, was für *F7c* auch der Fall ist. Bei *PAI* hingegen scheint das um ein Bp längere Allel mit geringfügig höherer Effizienz amplifiziert zu werden. Wahrscheinlichste Ursache für die geringfügige grö-

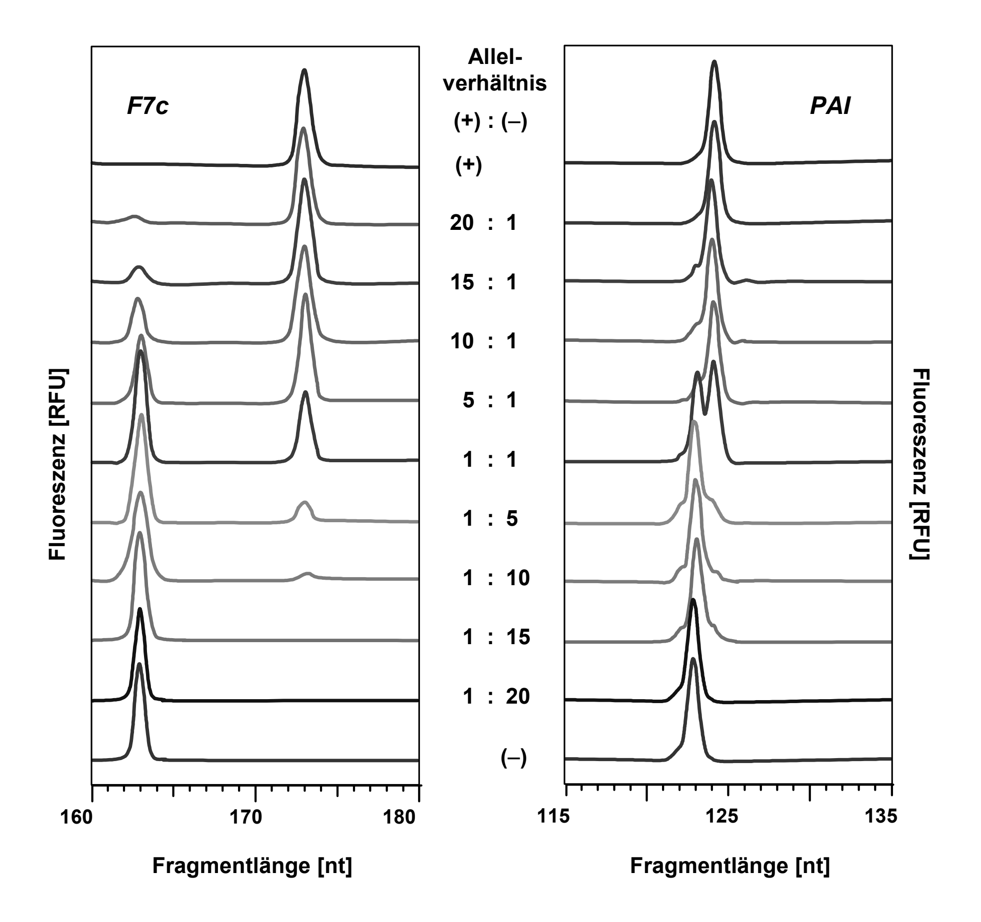


Abb. 3-51 Elektropherogramme der *F7c*- und *PAI*-spezifischen PCR-Produkte von Mischproben. Die FAM-markierten PCR-Produkte haben eine Länge von 163 Bp (–) bzw. 173 Bp (+) für *F7c* und 123 Bp (–) bzw. 124 Bp (+) für *PAI*.

ßere Menge des längeren PCR-Produkts ist ein kleiner Fehler (<10 %) bei der Konzentrationsbestimmung der homozygoten Proben. Zur Auswertung werden die gemessenen Peakflächenverhältnisse daher auf den Wert der Probe mit einem äquimolaren Mischungsverhältnis (1:1) normiert. Der nutzbare dynamische Bereich der Mischungsverhältnisse ergibt sich aus der Detektierbarkeit eindeutiger Peaks in den Elektropherogrammen.

Da in klinischen Studien in den meisten Fällen STRs zur Chimärenanalyse genutzt werden, wurde zusätzlich auch der gut charakterisierte $d(AAAAT)_n$-STR im *p53*-Gen [205,271] auf die gleiche Weise untersucht. Die Amplifikation eines den STR enthaltenden Abschnitts erfolgte mit dem Primerpaar VNTR-1b/2-HEX, welches je nach Anzahl der repetitiven Elemente HEX-markierte PCR-Produkte mit einer Länge zwischen 104 Bp bis 124 Bp liefert.

Für die Bestimmung des dynamischen Bereichs wurden hier die beiden kürzesten Allele verwendet, deren PCR-Produkte 104 Bp bzw. 109 Bp lang sind. Die Ergebnisse der Auswertungen der Peakflächenverhältnisse aus den Elektropherogrammen dieser Untersuchungen sind in **Abb. 3-52** zusammengefasst.

Während der dynamische Bereich für die *F7c*-Zielsequenz sowie für den STR zwischen 1:10 und 1:20 liegt, ist der dynamische Bereich für die *PAI*-Zielsequenz kleiner als 1:5, da in den Elektropherogrammen ausschließlich bei der Probe mit dem äquimolaren Mischungsverhältnis der beiden Allele zwei definierte Peaks erkennbar sind (siehe **Abb. 3-51**).

Die Abhängigkeit der Peakflächenverhältnisse (+):(–) vom eingesetzten Allelverhältnis (+):(–) lässt sich am besten durch eine Exponentialfunktion mit einem Exponenten kleiner eins beschreiben. Die Korrelationskoeffizienten für die in **Abb. 3-52** gezeigten Beispiele betragen 0.998 für *F7c* und 0.999 für den STR, während sie bei einem nach dem Modell bei identischen Amplifikationseffizienzen für beide Allele geforderten linearen Zusammenhang nur 0.935 bzw. 0.986 betragen (**Abb. 3-52**). Die Grenzwerte der Wurzelfunktion entsprechen den Grenzwerten der Allelverhältnisse (0 und $+\infty$), doch ihre Steigung nimmt

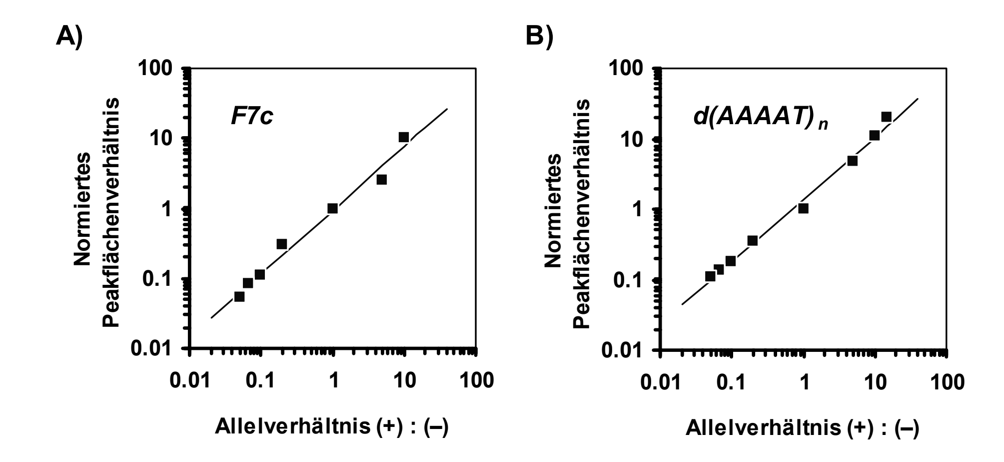


Abb. 3-52 Ergebnisse der kapillarelektrophoretischen Analyse von Mischproben. **A)** Ergebnisse für *F7c*. Regressionsfunktion: $y=0.9434 \cdot x^{0.908}$, $R^2 = 0.935$. **B)** Ergebnisse für *d(AAAAT)n*. Regressionsfunktion: $y=1.3841 \cdot x^{0.8821}$, $R^2 = 0.986$.

bei zunehmendem Überschuss des längeren Allels (+) ab, wodurch sich die Abweichung zur Achsenhalbierenden (die Peakflächenverhältnisse entsprechen hier den Allelverhältnissen) vergrößert. Dieses Verhalten ist dann zu erwarten, wenn das längere Allel (+) mit geringerer Effizienz amplifiziert wird, als das kürzere (–).

3.4.3.4 Analyse von Mischproben durch allelspezifische Real-time-PCR

Die *F7c*-allelspezifischen Primer für die RT-PCR sind so gewählt, dass die drei 3′-terminalen Basen des Primers R1 nicht zur (+)-Sequenz passen und die vier 3′-terminalen Basen von R2 nicht zur (–)-Sequenz passen, wenn der 5′-Terminus korrekt mit dem entsprechenden Templatstrang hybridisiert ist (**Abb. 3-49**). Obwohl hierdurch bedingt die 3′-Enden der Primer auf dem jeweils zu diskriminierenden Templat sehr instabil sind, werden sie unter den Versuchsbedingungen auch auf dem „falschen" Templat verlängert. Die Frequenz dieser Ereignisse ist beim Primer R2 geringer als beim Primer R1, obwohl R1 nur un-

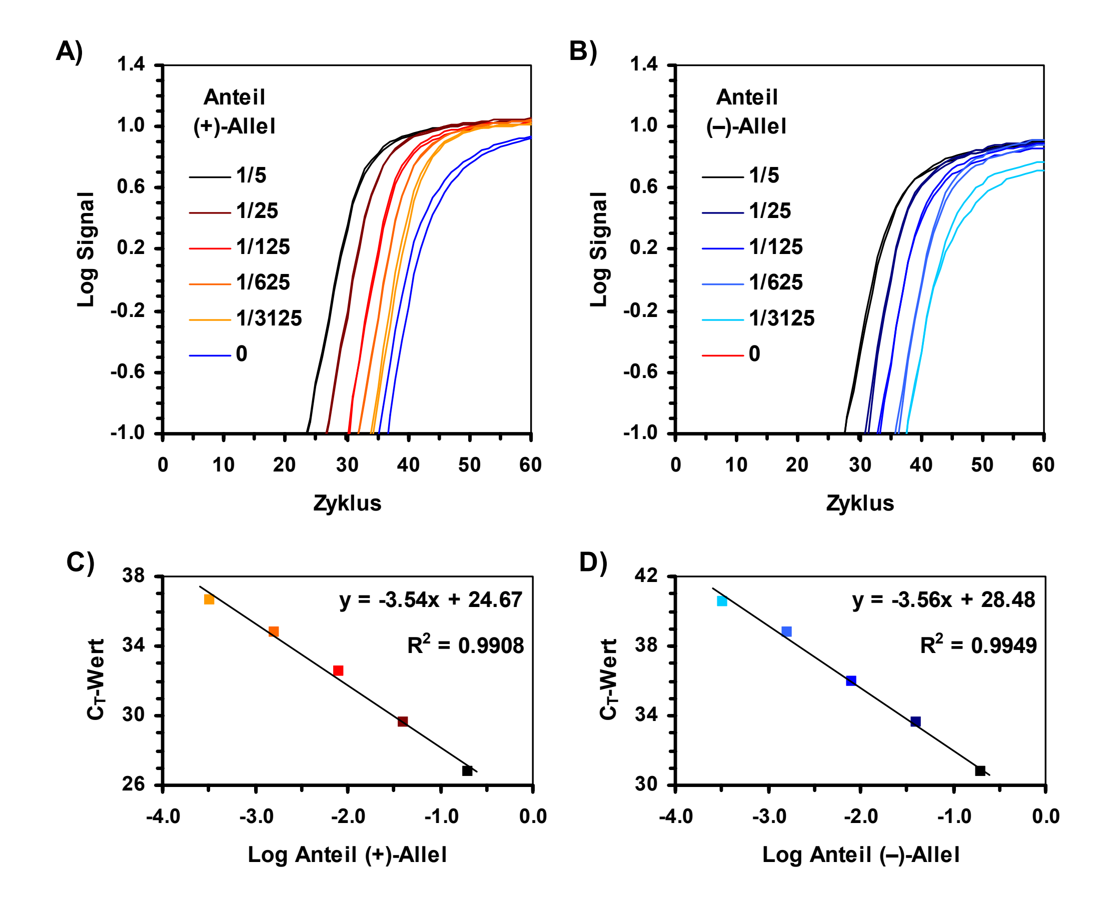


Abb. 3-53 Chimärenanalyse mittels *F7c*-allelspezifischer RT-PCR. A) und **B)** Amplifikationskurven für die Detektion des (+)-Allels bzw. des (–)-Allels. Die Negativkontrolle in B) zeigt keine Amplifikation. Templat: 20 ng humane genomische DNA, Detektion mit SYBR-Green I (Signalkurven von Kanal 1), Schwellenwert zur Bestimmung der C_T-Werte: jeweils 0.5. **C)** und **D)** Kalibriergeraden für die Quantifizierung des (+)-Allels bzw. des (–)-Allels aus den C_T-Werten der Kurven in A) bzw. B).

ter Ausbildung einer 9 Bp langen Schleife nichtgepaarter Nukleotide des (+)-Templatstrangs verlängert werden kann. In diesem Falle entsteht ein 138 Bp langes PCR-Produkt, was durch eine Schmelzkurvenanalyse und PAGE bestätigt wurde. Die letzten beiden Nukleotide am 3'-Ende von R2 sind hingegen komplementär zur Sequenz des (–)-Templats, so dass hier nur die folgenden 4 Nukleotide von R2 ungepaart bleiben, wenn dieser Primer auf dem (–)-Templatstrang verlängert wird. Das in diesem Falle entstehende Produkt hat dann ebenfalls eine Länge von 138 Bp.

Die Ergebnisse der Chimärenanalyse mittels RT-PCR sind am Beispiel des 10 Bp-Polymorphismus im *F7c*-Promotor in **Abb. 3-53** gezeigt. Jede Probe wurde zweifach gemessen, um die hohe Reproduzierbarkeit dieses Systems zu demonstrieren. Bis auf die Negativkontrolle für die Quantifizierung des (+)-Allels sind die Kurvenpaare in **Abb. 3-53 A, B** nahezu deckungsgleich.

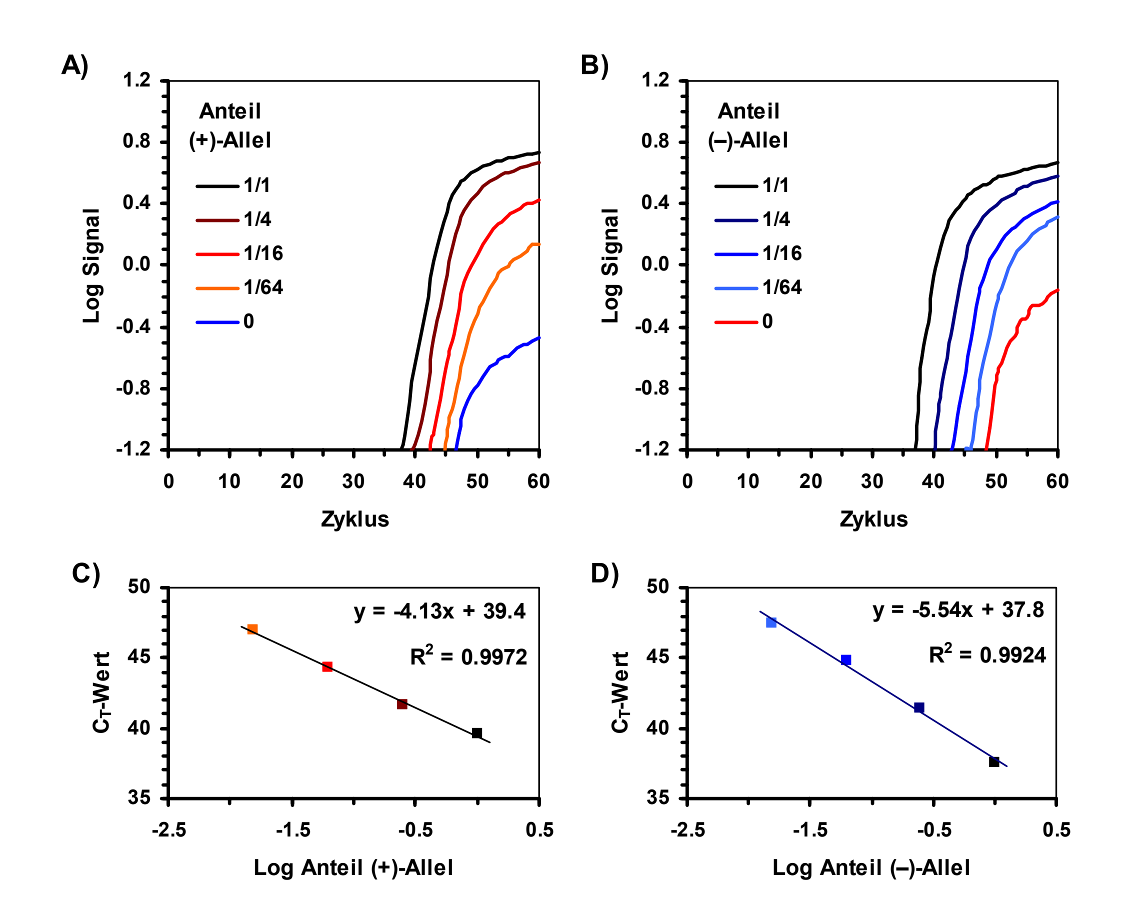


Abb. 3-54 Chimärenanalyse mittels *PAI*-allelspezifischer RT-PCR. A) und **B)** Amplifikationskurven für die Detektion des (+)-Allels (Primer PAI-H2/R2) bzw. des (–)-Allels (Primer PAI-H1/R2). Templat: 20 ng humane genomische DNA, Detektion mit SYBR-Green I (Signalkurven von Kanal 1), Schwellenwert zur Bestimmung der C_T-Werte: 0.18 für A) und 0.2 für B). **C)** und **D)** Kalibriergeraden für die Quantifizierung des (+)-Allels bzw. des (–)-Allels aus den C_T-Werten der Kurven in A) bzw. B).

Die Ergebnisse zeigen einen linearen Zusammenhang zwischen den C_T-Werten und dem Logarithmus der initialen Menge des nachzuweisenden Allels bis zu einem mehr als 3000-fachen Überschuss des zu diskriminierenden Allels, womit dieses Verfahren bei hoher Präzision einen dynamischen Bereich von mehr als 1:3000 aufweist. In den Negativkontrollen entstand ausschließlich das für die Primer spezifische Produkt, was durch Schmelzkurvenanalysen und PAGE charakterisiert wurde: Hier erfolgte die Verlängerung von auf dem „falschen" Templat hybridisierten Primern. Die entstandenen Produkte, welche die Sequenz des „richtigen" Templats aufweisen, wurden anschließend amplifiziert.

Für den 4G/5G-Polymorphismus im *PAI*-Promotor wurde in gleicher Weise wie für den *F7c*-Polymorphismus ein dynamischer Bereich von 1:64 bestimmt (**Abb. 3-54**). Dieser Wert ist wie erwartet kleiner als der für *F7c* erreichte Wert, jedoch mehr als dreimal größer als der mit kompetitiven Amplifikationstechniken erreichbare Wert von ca. 1:20 und sogar mehr als zehnmal größer als der in der kapillarelektrophoretischen Analyse erreichte Wert von maximal 1:5.

Wie bei *F7c* auch, handelt es sich bei den in den Negativkontrollen entstandenen PCR-Produkten ausschließlich um die für die Primer spezifischen Produkte und nicht um unspezifische Produkte wie z. B. Primerdimere, die bei einer Detektion mit SYBR-Green I ähnliche Amplifikationskurven liefern würden. Die entstandenen Produkte wurden im Anschluss an die PCR durch Schmelzkurvenanalysen und PAGE charakterisiert.

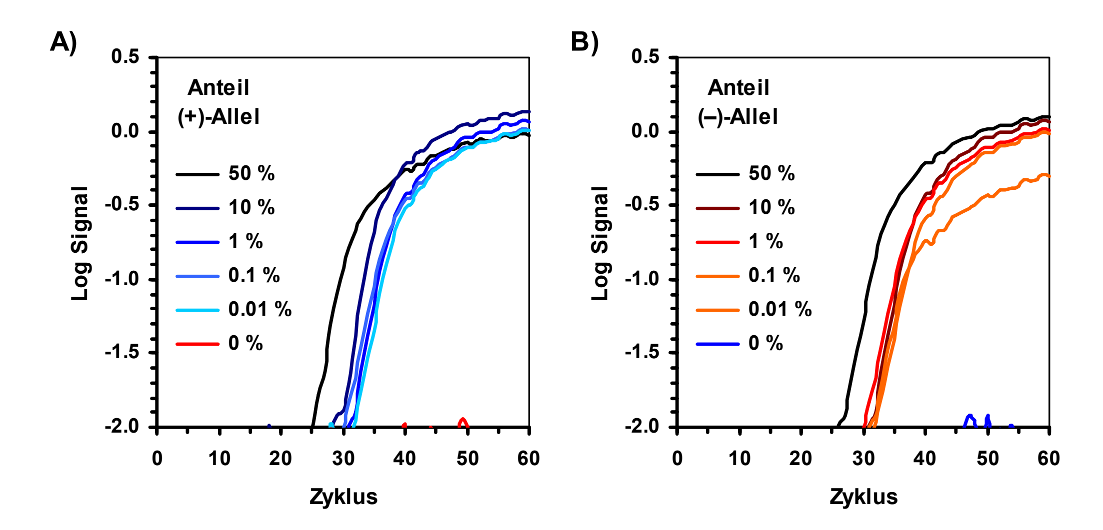


Abb. 3-55 Detektion von *PAI*-Allelen in Mischproben. A) Nachweis des (+)-Allels bzw. **B)** des (–)-Allels. Die Proben enthalten jeweils 30 ng humane genomische DNA (entsprechend 10 000 haploiden Genomkopien) mit den angegebenen Anteilen des nachzuweisenden Allels. Die *Taq* DNA-Polymerase wurde vor der PCR mit TaqStart-Antikörper inkubiert und während der fünfminütigen initialen Denaturierungsphase aktiviert. Die Amplifikation erfolgte abweichend vom Standardprotokoll mit einer Hybridisierungstemperatur von 59 °C und einer verlängerten Denaturierungsphase von 5 s je Zyklus. Die Detektion der PCR-Produkte erfolgte mit SYBR-Green I.

Die Diskriminierung des im Überschuss befindlichen Allels kann durch stringentere Versuchsbedingungen beträchtlich erhöht werden: Eine höhere Hybridisierungstemperatur (59 °C statt 55 °C) sowie die Anwendung einer *Hot-start*-Technik (**0**) ermöglichte die vollständige Diskriminierung eines 10 000-fachen Überschusses des zu diskriminierenden Allels bzw. den eindeutigen Nachweis einer einzigen Templatkopie des zu amplifizierenden Allels in einer Mischung mit dem 10 000-fachen Überschuss des zu diskriminierenden Allels (**Abb. 3-45**). Unter diesen Bedingungen geht jedoch der lineare Zusammenhang zwischen den Logarithmen der Templatkopienzahl und den C_T-Werten verloren, so dass sich diese Variante zwar zum hochempfindlichen und selektiven Nachweis eines bestimmten Allels, nicht jedoch zu dessen Quantifizierung einsetzen lässt.

Die Ursache hierfür ist nicht geklärt. Innerhalb der ersten Amplifikationszyklen scheint unter den gegebenen Bedingungen die Wahrscheinlichkeit, dass ein Primer auf einem Templatmolekül überhaupt verlängert wird, sehr gering zu sein, so dass die Verlängerung des Primers auf einem zu diskriminierenden Templatstrang praktisch nicht erfolgt. Wenn jedoch einmal eine Verlängerung stattgefunden hat und PCR-Produkt akkumuliert, steigt die Amplifikationseffizienz in den späteren Zyklen schließlich auf Werte nahe Zwei. Der Einfluss statistischer Schwankungen auf Zeitpunkt und Anzahl der ersten Verlängerungsereignisse ist jedoch gerade bei geringen Templatkopienzahlen im Reaktionsansatz nicht mehr vernachlässigbar und könnte der Grund für die nicht mehr vorhandene Korrelation der Templatkopienzahlen und der C_T-Werte sein.

Zwar handelt es sich bei dem 4G/5G-Polymorphismus im *PAI-1*-Promotor um einen Längen- bzw. Insertions-/Deletions-Polymorphismus mit einer Insertion von einem G:C-Basenpaar, jedoch stellt sich diese Variation dem hier beschriebenen Verfahren wie ein Basenaustausch dar, d. h. wie ein *single nucleotide polymorphism* (SNP). Der diskriminierende Primer kann die beiden Allele nur anhand seiner 3'-terminalen Base unterscheiden. Im Falle der Hybridisierung mit dem „falschen" Templatstrang liegt am 3'-Ende entweder statt einer A:T-Paarung eine A:G-Fehlpaarung (PAI-H1 × (+)-Allel) oder statt einer G:C-Paarung eine G:T-Fehlpaarung vor (PAI-H2 × (–)-Allel). Der 4G/5G-Polymorphismus ist also für die hier vorgestellte Form einer allelspezifischen PCR nicht von einem G/A (bzw. C/T) SNP unterscheidbar. Die an sich schon hohe Anforderung an das Diskriminierungspotenzial der Primer, die hier die Sequenzabweichung an nur einer einzigen Position unterscheiden müssen, ist in dem hier gezeigten Beispiel nochmals höher, da sich die polymorphe Stelle zudem unmittelbar neben einer thermodynamisch stabilen G:C-Klammer befindet.

3.4.3.5 Zusammenhang von dynamischem Bereich, ΔC_T-Werten und der Amplifikationseffizienz

Der dynamische Bereich dieses Verfahrens für die Quantifizierung eines Allels in Mischproben lässt sich anhand der C_T-Werte der Amplifikationskurven für die nachzuweisende Zielsequenz (A) und die zu diskriminierende Sequenz (B) berechnen, deren Differenz $C_{T,B} - C_{T,A}$ als $\Delta C_{T,BA}$ bezeichnet wird. Der Wert von $C_{T,A}$ ist nach **Gl. 1.8** abhängig von der Templatkopienzahl $N_{0,A}$. Nahezu jede Templatkopie (und nachfolgend jedes Produkt) wird in den

PCR-Zyklen kopiert. Wird statt des Templats A die gleiche Menge $N_{0,B} = N_{0,A} = N_0$ des Templats B in die Reaktion eingesetzt, ist im Idealfall keine Amplifikation zu erwarten, da die Template nicht kopiert werden. Tatsächlich jedoch ist die Wahrscheinlichkeit der Verlängerung eines mit dem „falschen" Templat B hybridisierten Primers (und die dann in den nächsten Zyklen folgende Amplifikation der PCR-Produkte) in den meisten Fällen größer als Null, so dass auch hier eine Amplifikationskurve mit endlichem Wert für $C_{T,B}$ aufgezeichnet wird. Von den eingesetzten N_0 Kopien des Templats B werden jedoch nicht alle im ersten Zyklus kopiert. Je besser der allelspezifische Primer die Templatsequenz B diskriminiert, desto geringer ist die Wahrscheinlichkeit p eines Kopierereignisses und mithin der Anteil kopierter Templatmoleküle, der sich zu $p{\cdot}N_0$ ergibt (mit $0 \leq p \leq 1$). Das Produkt, welches durch diesen Kopiervorgang entsteht, hat eine Primerbindungsstelle, die vollständig komplementär zur Primersequenz und somit identisch mit dem A-Produkt ist. Es wird folglich mit der gleichen Effizienz ε amplifiziert wie PCR-Produkte des Templats A. Die C_T-Werte der Amplifikationskurven für einen Schwellenwert T sind $C_{T,A} = \log_\varepsilon(T{\cdot}N_0^{-1})$ und $C_{T,B} = \log_\varepsilon(p{\cdot}T{\cdot}N_0^{-1}) = \log_\varepsilon(p) + \log_\varepsilon(T{\cdot}N_0^{-1})$. Die Differenz $\Delta C_{T,BA} = C_{T,B} - C_{T,A} = \log_\varepsilon(p)$ ist damit lediglich abhängig von der Wahrscheinlichkeit p, die als Diskriminierungswert bezeichnet werden kann, und der Amplifikationseffizienz ε, die anhand der Steigung einer Standardkurve bestimmt werden kann (siehe **Gl. 1.6** und **Gl. 1.8**). Für den Diskriminierungswert p ergibt sich damit aus den für die Template A und B bestimmten C_T-Werten $p = \varepsilon^{\Delta C_{T,BA}}$.

3.4.3.6 Betrachtung zur Anwendung der SNP-Analytik für die Chimärenanalyse

Chimäre Organismen, Organe oder Gewebe bestehen aus Zellen von ursprünglich verschiedenen Individuen. Die Chimärenanalyse umfasst die Quantifizierung der Zellen in Abhängigkeit von ihrem Ursprung, wozu charakteristische Unterschiede genutzt werden, die letztlich auf Unterschiede im genetischen Material der Zellen zurückzuführen sind. Die zu erwartenden Unterschiede sind jedoch umso kleiner, je enger die (ursprünglichen) Individuen miteinander verwandt sind. Ein hoher Verwandtschaftsgrad ist jedoch häufig eine Voraussetzung für die Überlebensfähigkeit der Chimäre, weil es sonst z. B. zu extremen immunologischen Unverträglichkeitsreaktionen kommen kann. So werden beispielsweise zur heterologen Knochenmarktransplantation bei leukämiekranken Personen, deren Ergebnis meist ein chimäres hämatopoetisches System ist, bevorzugt sehr enge Verwandte als Spender herangezogen. Damit ist die Wahrscheinlichkeit, ausgeprägte genetische Unterschiede im Erbgut von Spender und Empfänger zu finden, wie sie zur Chimärenanalytik verwendet werden könnten, sehr gering.

Ein diesbezüglich gutes Ziel für eine Chimärenanalytik stellen hier die SNPs dar, die in sehr großer Zahl im Genom vorkommen.[58,59] Alleine durch ihre hohe Dichte im Genom ist die Wahrscheinlichkeit groß, auch in nahe verwandten Individuen unterschiedliche Ausprägungen von SNPs zu finden. Wie groß diese Wahrscheinlichkeit ist, hängt von der Allelfrequenz sowie der untersuchten Anzahl von SNPs ab. Da die meisten SNPs biallelisch sind, d. h. in zwei möglichen Ausprägungen (A und B) vorkommen, können die Allelfrequenzen durch a und b beschrieben werden, wobei $b = 1 - a$. Die Wahrscheinlichkeit p, eine brauch-

bare Kombination von Ausprägungen in zwei Individuen zu finden, entspricht der Summe der Wahrscheinlichkeiten für alle möglichen Fälle, dass eine Ausprägung exklusiv nur bei einem Individuum auftritt:

$$P = q(AA/BB) + q(AB/AA) + q(AB/BB) + q(BB/AA)$$

wobei AA, AB und BB den möglichen Genotypen entsprechen. Die Wahrscheinlichkeiten, einen bestimmten Genotyp zu finden, betragen gemäß der binomischen Verteilung

$$q(AA) = a^2,\ q(AB) = 2 \cdot a \cdot b \text{ und } q(BB) = b^2$$

Für p folgt damit

$$p = a^2 \cdot b^2 + 2 \cdot a \cdot b \cdot a^2 + 2 \cdot a \cdot b \cdot b^2 + b^2 \cdot a^2$$

$$p = a^2 \cdot b^2 + 2 \cdot a \cdot b \cdot a^2 + 2 \cdot a \cdot b \cdot b^3 + b^2 \cdot a^2 = 2 \cdot \left(a^3 b + a^2 b^2 + ab^3\right) \qquad \textbf{Gl. 3.34}$$

Nach Einsetzten von $b = 1-a$ in **Gl. 3.34** und Umformung ergibt sich

$$p = 2 \cdot \left(a - 2a^2 + 2a^3 - a^4\right) \qquad \textbf{Gl. 3.35}$$

Diese Funktion besitzt ein Maximum bei $a = 0.5$, wie sich durch Lösen von $\mathrm{d}P/\mathrm{d}a = 0$ berechnen lässt. Die Lösung besagt, dass bei einer Allelfrequenz von $a = b = 0.5$ des untersuchten SNP die Wahrscheinlichkeit p maximal ist, eine geeignete Kombination von Genotypen bei zwei Individuen zu finden. Setzt man den Wert von $a = 0.5$ in **Gl. 3.35** ein, vereinfacht sich diese zu

$$p = 2 \cdot \left(2a^3 - a^4\right) \qquad \textbf{Gl. 3.36}$$

Die inverse kumulative Wahrscheinlichkeit P_c, also die Wahrscheinlichkeit, mindestens eine brauchbare Kombination in einem Satz von n betrachteten SNPs mit einer Allelfrequenz von je 0.5 zu finden, ist gleich der inversen Kumulativen von p: $1-P_c = (1-p)^n = (1 - 2 \cdot (2a^3 - a^4))^n$. Aufgelöst nach n ergibt sich

$$n = \frac{\ln\left(1 - P_c\right)}{\ln\left(1 - 4 \cdot a^3 + 2 \cdot a^4\right)} \qquad \textbf{Gl. 3.37}$$

Anhand dieser Gleichung lässt sich nun die benötigte Anzahl n von zu untersuchenden SNPs berechnen, um mit der gewünschten Wahrscheinlichkeit P_c eine brauchbare Kombination zu finden. Möchte man z. B. mit mindestens 99%iger Wahrscheinlichkeit anhand der untersuchten SNPs eine Chimärenanalyse durchführen können, muss der Satz 10 SNPs umfassen ($n = 9.8$ für $P_c = 0.99$; $a = 0.5$).

4 GESAMTDISKUSSION

Vielwisserei verleiht nicht Verstand;
sonst hätte sie dem Hesiod und Pythagoras solchen verliehen
und ebenso dem Xenophanes und Hekatäos.

Heraklit

DNA-analytische Verfahren sind von großer Bedeutung in der molekularen Biologie und Medizin. Stand in den letzten Dekaden noch die Sequenzaufklärung von Genen und Genomen im Vordergrund, so beginnen wir heute, die individuellen Unterschiede in den Genomen zu untersuchen und deren Einfluss auf Entwicklungs- und Krankheitsprozesse zu verstehen. Die Unterschiede betreffen dabei nicht nur die Sequenzen an sich, wie z. B. die Identifizierung von Mutationen oder *single nucleotide polymorphisms* (SNPs), sondern auch die Anzahl bestimmter Sequenzen bzw. Gene im Genom, insbesondere die tumorrelevanter Gene.[272] Dabei ist die PCR heute die wichtigste methodische Grundlage zur Untersuchung von Nukleinsäuren, da sich mit ihrer Hilfe interessierende Sequenzabschnitte aus kaum direkt nachweisbaren Mengen von DNA-Molekülen milliardenfach kopieren und so einer Analyse zugänglich machen lassen. Dieser Amplifikationsprozess selbst kann bei RT-PCR-Verfahren durch die Anwendung spezieller Fluoreszenztechniken sichtbar gemacht werden.

Die vorliegende Arbeit wurde zu einem Zeitpunkt begonnen, als die ersten RT-PCR-Geräte in Großserien verkauft wurden. Das Ziel der Arbeit war zunächst die Entwicklung von Methoden zur hochgenauen Quantifizierung tumorrelevanter Gene, um z. B. den Deletionsgrad von Tumorsuppressorgenen in Tumorproben zu bestimmen (gezeigt am Bespiel von *p53* in **3.3.3**). Hierzu ist ein großer dynamischer Bereich für die Quantifizierung, wie ihn die RT-PCR bietet, nicht notwendig, wohl aber eine sehr hohe Genauigkeit. Im Falle der Bestimmung von Deletionsgraden von Tumorsuppressorgenen, wo im Idealfall jede in der Probe analysierte (Tumor-)Zelle nur noch eines der beiden Allele des Gens besitzt, entspricht der Messbereich gerade einem Faktor von zwei. Weiterhin muss, zur Normierung auf die Probengröße, zusätzlich ein Referenzgen quantifiziert werden, was ebenfalls mit einem Fehler behaftet ist, wodurch die Auflösung des normierten Ergebnisses nochmals verringert wird. Ein zentraler Aspekt dieser Arbeit bestand daher darin, durch die Entwicklung entsprechender Verfahren zur Durchführung und Auswertung möglichst hohe Genauigkeiten zu erzielen. Dies wurde durch die umfassende Analyse aller relevanten biochemischen Parameter und die Entwicklung neuer Verfahren zur Auswertung der Messdaten erreicht.

4.1 Anwendungsgebiete der Real-time-PCR

Die Kenntnis der Reaktionsverläufe kann genutzt werden, um schnell und einfach die initiale Menge des Templats zu bestimmen.[22,23] Anhand der Aufzeichnung der Fluoreszenzinten-

sität in Abhängigkeit von der Temperatur bei sogenannten Schmelzkurvenanalysen lassen sich die amplifizierten Sequenzabschnitte hinsichtlich Länge oder Sequenzvariationen weiter charakterisieren.[97,219] Eine nachträgliche, externe Analyse der PCR-Produkte wie z. B. mittels Gelelektrophorese ist nicht mehr notwendig, die Reaktionsgefäße brauchen im Anschluss an die PCR nicht mehr geöffnet zu werden. Dies mindert die Gefahr von Kontaminationen der Arbeitsplätze mit PCR-Produkten erheblich.

Die entscheidenden Vorteile der RT-PCR gegenüber anderen Verfahren PCR-basierter Quantifizierungen von Nukleinsäuren sind der extrem große dynamische Bereich (über mehr als 8 Größenordnungen)[23] und der hohe Grad an Verlässlichkeit der Ergebnisse, da nicht nur die Produktmenge einmalig nach der PCR analysiert wird, sondern die Kinetik der gesamten Reaktion sichtbar ist. Abweichende Amplifikationseffizienzen in den individuellen Reaktionsansätzen lassen sich somit leicht identifizieren. Die RT-PCR ist damit hervorragend geeignet, schnell und sicher die Menge bestimmter Nukleinsäuren, z. B. die DNA oder RNA von Viren[273] oder Transkriptmengen[18,39,274] in biologischem Probenmaterial zu ermitteln und bildet die methodische Grundlage einfacher, schneller und verlässlicher Bestimmungen der Titer von Krankheitserregern in Nahrungsmitteln,[275-279] Blut oder anderen Organen bzw. Körperflüssigkeiten, für Expressionsstudien,[z. B. 280] die Bestimmung von Allelungleichgewichten[253] sowie die direkte Quantifizierung von Gendeletionen und Amplifikationen[281,282] und nicht zuletzt für die Detektion und Verfolgung von Minimalresidualkrankheiten (*minimal residual diseases*) bei Leukämien anhand leukämiespezifischer Fusionstranskripte.[283-287] Ein weiteres, zunehmend wichtiges Anwendungsfeld der RT-PCR ist die Bestätigung und Kontrolle der Daten von *Microarray*-Analysen.[288-290]

4.2 Die Marktentwicklung von Real-time-PCR-Systemen

Die ersten Real-time-Thermocycler wurden 1997 von der Firma Applied Biosystems vertrieben. Dem heute noch marktführenden Gerät ABI 7700 folgte wenig später der LightCycler der Firma Roche, welche die Lizenz für diese Technologie von Idaho Technology erwarb. Zwar können mit dem LightCycler nur 32 Proben parallel analysiert werden – im Gegensatz zum ABI 7700, der 96 parallele Analysen ermöglicht – doch sind beim LightCycler durch die besondere Konstruktion extrem kurze Zykluszeiten möglich, so dass eine Analyse in nur 15 bis 20 Minuten abgeschlossen ist (**2.6.2**), während das Gerät von Applied Biosystems 2 bis 4 Stunden benötigt. Schließlich entwickelte die Firma Bio-Rad mit dem iCycler ein eigenes RT-PCR-Gerät, dessen wesentliche Neuerung ein Fluoreszenzdetektionsmodul darstellt, welches einem herkömmlichen Thermocyler aufgesetzt ist. Inzwischen werden RT-PCR-Systeme u. a. auch von den Firmen MJ Research, Stratagene, Thermo Hybaid, Corbett Research und Cepheid angeboten. Weitere Firmen werden folgen. So steht bei Eppendorff ein weiteres System unmittelbar vor der Markteinführung.

Die Investitionen der Unternehmen in die Entwicklung von RT-PCR-Systemen sind durchaus gewinnträchtig. Dem Frost & Sullivan Report B114 von 2002 zufolge hat der Markt für solche Systeme in den nächsten Jahren noch ein beträchtliches Wachstumspotenzial:[291] Im

Jahre 2002 betrug der Umsatz für RT-PCR-Systeme in den USA und Europa etwa 310 Millionen US-Dollar (Europa: 120 Millionen US-Dollar), wobei die drei Marktführer Applied Biosystems, Roche und Bio-Rad durch den frühen Einstieg und die Nutzung bestehender Vertriebsstrukturen mit einem Marktanteil von 85 % vorherrschen. Alleine in Europa wird bis 2006 ein Umsatzanstieg auf über 350 Millionen US-Dollar prognostiziert, während beim Umsatz von konventionellen Thermocyclern im gleichen Zeitraum von etwa 80 Millionen nur ein Anstieg auf 130 Millionen US-Dollar erwartet wird. Die bedeutendsten Abnehmer für RT-PCR-Geräte sind die klinische Diagnostik und die Lebensmittelprüfung, zunehmend wichtig werden die Lebensmittelindustrie sowie die Gerichtsmedizin. In diesen Gebieten ersetzt die RT-PCR die vergleichsweise zeitintensiven Immunotests (z. B. ELISA, *enzyme-linked immunosorbent assay*), die konventionellen PCR-basierten Verfahren noch vorgezogen werden. Die RT-PCR ist jedoch auf Grund der hohen Reproduzierbarkeit der Ergebnisse durch die Kenntnis der Amplifikationskurven, also des vollständigen Reaktionsverlaufs, gegenüber den konventionellen PCR-basierten Verfahren erheblich an Zuverlässigkeit gewonnen.[292]

4.3 Möglichkeiten und Grenzen der Real-time-PCR

Mit der weiten Verbreitung der Real-time-Systeme in den letzten Jahren stieg auch die Zahl der Publikationen über RT-PCR exponentiell an. Die neuen Techniken fanden sehr schnell Anwendung in der medizinischen Forschung und Diagnostik,[272,273,292] obwohl bisher noch kaum Veröffentlichungen über systematische Untersuchungen ihrer Grundlagen, Möglichkeiten und Grenzen vorliegen. So werden beispielsweise häufig hohe Amplifikationseffizienzen als Qualitätskriterium angegeben, ohne jedoch zu zeigen, warum hohe Effizienzen besser sein sollten als niedrigere. Wie in **3.1.2** gezeigt, steigt die Präzision der Bestimmung der C_T-Werte tatsächlich leicht mit steigender Effizienz, die Präzision der Bestimmung der eigentlich interessierenden Templatkonzentrationen nimmt jedoch sogar leicht ab. Insgesamt ist der Einfluss der Effizienz auf die Genauigkeit der Ergebnisse vernachlässigbar gering, nicht allerdings auf die Sensitivität. Bei geringerer Effizienz steigt die benötigte Anzahl an Zyklen zum Nachweis der Zielsequenz schnell an und mithin auch die Wahrscheinlichkeit der Bildung von unspezifischen Nebenprodukten, die in den weiteren Zyklen mit hoher Effizienz amplifiziert werden können und so die Amplifikation der Zielsequenz durch Kompetition weiter verschlechtern oder gar verhindern.

Die erreichbare Genauigkeit von Quantifizierungen mittels der RT-PCR wird nicht selten überschätzt. Sie ist für die derzeit üblichen Verfahren, bedingt durch den exponentiellen Zusammenhang der fehlerbehafteten Messgröße (C_T-Wert) und der daraus berechneten Templatmenge, deutlich geringer als die konventioneller Verfahren, insbesondere als die der kompetitiven PCR.[17,19,20] So gibt Roche Diagnostics an, dass die praktische Auflösung einer RT-PCR etwa einem Faktor von 2 entspricht, also einem Variationskoeffizienten von 50-100 % (dies entspricht einem Fehler in der Bestimmung des C_T-Wertes von weniger als

±0.5 Zyklen), während mit der kompetitiven PCR Variationskoeffizienten von etwa 10 % erreichbar sind.[205]

Ein weiteres, häufiges Vorurteil ist die Überlegenheit von sequenzspezifischen Sonden gegenüber dsDNA-sensitiven Farbstoffen wie SYBR Green I hinsichtlich der erreichbaren Genauigkeit, insbesondere bei geringen Templatmengen.[27,293,294] Neben den grundsätzlichen, rein statistischen Schwankungen der Templatmengen in den einzelnen Reaktionsansätzen, dem sogenannten Poisson-Fehler,[213,295,296] die durch entsprechende statistische Verfahren korrigiert werden können, führen die Bildung und Amplifikation unspezifischer Produkte in den späten Phasen der PCR zu systematischen Fehlern, da hier die Amplifikationseffizienz für die Zielsequenz verringert wird. Dieser Zusammenhang gilt selbstverständlich auch dann, wenn diese unspezifischen Produkte z. B. durch die Verwendung sequenzspezifischer Sonden nicht detektiert werden. In diesem Falle erhält man eine systematische Unterbestimmung der Templatmenge. Die Größe des Fehlers lässt sich nicht feststellen, ohne die Kinetik der Bildung der unspezifischen Produkte zu kennen. Die Vorteile einer sequenzspezifischen Detektion liegen vielmehr in der Möglichkeit des Nachweises bestimmter Template, die in nur sehr geringen Konzentrationen im Reaktionsansatz vorliegen und neben anderen Produkten (z. B. Primer-Dimeren) amplifiziert werden, sowie der Möglichkeit, Sequenzvarianten zu unterscheiden und mehrere unterschiedliche Zielsequenzen im selben Reaktionsansatz gemeinsam analysieren zu können.

Den größten Einfluss auf die erreichbare Genauigkeit bei Quantifizierungen haben gerätespezifische Parameter wie die Temperaturhomogenität sowie die Sensitivität und die Genauigkeit der Fluoreszenzmessung (**3.1.2**): Temperaturinhomogenitäten im Probenraum des Geräts können zu unterschiedlichen Amplifikationseffizienzen in Abhängigkeit der Probenposition im Gerät führen. Besonders anfällig für derartige Effekte erscheinen Zielsequenzen zu sein, die sich in der Nähe von hochschmelzenden Regionen genomischer Template befinden oder solche Regionen sogar selbst enthalten. Diese Bereiche werden während der Denaturierungsphase mitunter nicht vollständig aufgeschmolzen (**3.1.4**). Hierdurch bedingt ist eine verlässliche Quantifizierung anhand externer Standards nicht möglich. Die Probleme können allerdings leicht vermieden werden, z. B. durch eine gründliche initiale Denaturierung der genomischen Template im kochenden Wasserbad. Im LightCycler wurde nach der Feststellung und Analyse derartiger Probleme vom Hersteller die Option eingerichtet, das Probenkarussell während der PCR kontinuierlich rotieren zu lassen, um den Einfluss positionsabhängiger Temperaturunterschiede im Gerät zu minimieren.

Die Sensitivität des Meßsystems begrenzt die Anzahl der Messpunkte mit erkennbarem exponentiellen Anstieg (eine höhere Sensitivität erlaubt die Messung exponentieller Produktzunahme schon in früheren Zyklen), die zur Ermittlung der C_T-Werte verwendet werden können. Der Messfehler wiederum begrenzt die erreichbare Genauigkeit bei einer gegebenen Anzahl verwendeter bzw. geeigneter Datenpunkte. Die effektive Sensitivität der Fluoreszenzmessung ist auch von der Produktausbeute abhängig. Wird mehr Produkt gebildet, also ein höherer Plateauwert erreicht, befinden sich im Allgemeinen auch mehr Datenpunkte der Amplifikationskurve in der exponentiellen Phase und der Einfluss der Plateau-

phase auf die C_T-Werte verringert sich (**3.1.3**). Einen nennenswerten Einfluss auf die Ausbeute und damit auf die effektive Sensitivität hat die Magnesiumionenkonzentration im Reaktionsansatz (**3.1.1**). Dies gilt auch bei Verwendung von Hybridisierungssonden, die mit steigender Magnesiumionenkonzentration während der PCR verstärkt hydrolysiert werden (**3.1.5**): Die Effekte der Sondenhydrolyse werden durch die höhere Gesamtausbeute mehr als ausgeglichen.

Die Genauigkeit von Quantifizierungen ist zudem abhängig von der Lage des Schwellenwerts, der zur Ermittlung der C_T-Werte verwendet wird. Während die Fehler der Fluoreszenzmessungen (das Signalrauschen) durch höhere Schwellenwerte verringert werden, ist der Einfluss der Plateauphase bei niedrigeren Schwellenwerten geringer. Zwischen den Extremwerten für die Lage des Schwellenwerts muss somit ein Wert existieren, bei welchem der dort nur moderate Einfluss beider Fehlerquellen in der Summe minimal ist. Dies ist tatsächlich der Fall, wie in dieser Arbeit anhand der Fehlerfunktion der Kalibriergeraden sowie der Streuung der C_T-Werte von Mehrfachansätzen erstmals experimentell gezeigt werden konnte (**3.2.1.5**).

4.4 Quantitative Analyse

Neben den technischen, geräte- und reaktionsbedingten Parametern hat die Auswertung der Messdaten, hier insbesondere die Berechnung der C_T-Werte, einen entscheidenden Einfluss auf die Genauigkeit der Ergebnisse von Quantifizierungen. Die von den Marktführern angebotenen Programme zur Auswertung von RT-PCR-Experimenten nutzen die *Fit-points*-Methode (siehe **3.2.1.1**) zur Berechnung der C_T-Werte, die mehrere Schwachstellen aufweist. Hervorzuheben sind hier die nicht durchgeführte Korrektur der Signalkurven um amplifikationsunabhängige Trends, welche die tatsächliche exponentielle Signalzunahme maskieren sowie die hohe Rauschempfindlichkeit der Regressionsgeraden in Abhängigkeit von den verwendeten *„fit points“*. Die Anzahl dieser Punkte ist bei den kommerziellen Programmen für alle Signalkurven eines Experiments identisch. Die Signalkurve mit der kürzesten erkennbaren exponentiellen Phase setzt daher die Obergrenze für die Zahl der verwendbaren *fit points*. Dieses Auswerteverfahren erfordert die Beurteilung der Signalkurven und die Identifizierung deren exponentieller Abschnitte durch den Experimentator, der auch die Lage des Schwellenwerts festlegen muss. Um die Analyse weiter zu automatisieren und die willkürliche Einflussnahme durch die manuellen Einstellungen des Experimentators zu verringern, führte Roche Diagnostics im Jahr 2000 die 2^{nd}-*Derivative-maximum*-Methode als eine neue Methode zur Bestimmung der C_T-Werte ein. Hier werden die C_T-Werte als das Maximum der zweiten Ableitung (dem Punkt der maximalen Krümmung) der geglätteten Signalkurven berechnet. Dieses Verfahren ist skalierungsinvariant, benötigt keinen Schwellenwert und der Einfluss amplifikationsunabhängiger Signaltrends auf den C_T-Wert ist deutlich geringer als bei der *Fit-points*-Methode, da die Trends im Bereich des Krümmungsmaximums der Signalkurven näherungsweise linear sind. Die 2^{nd}-*derivative-maximum*-Methode liefert in der Praxis jedoch nur wenig reproduzierbare Ergebnisse, weshalb Roche Diagnostics das Programm 2001 um einen sogenannten Plausibilitäts-Check

erweitert hat, der zumindest anzeigt, ob das berechnete Ergebnis möglicherweise falsch ist. Ein weiteres prinzipielles Problem dieser Methode besteht darin, dass die C_T-Werte nicht nur vom Plateauwert selbst abhängig sind, sondern auch von der Charakteristik des Übergangs von der exponentiellen Phase in die Plateauphase der Signalkurven. Aus diesen Gründen hat sich diese Methode in der Praxis nicht durchgesetzt.

Die in dieser Arbeit entwickelte Methode zur quantitativen Auswertung von RT-PCR-Daten korrigiert die Signalkurven um amplifikationsunabhängige Trends, identifiziert die exponentiellen Phasen und verwendet von jeder Signalkurve alle Datenpunkte innerhalb der erkannten exponentiellen Phase zur Berechnung der C_T-Werte. Es ist zwar nicht bekannt, wodurch die Trends der Signalkurven verursacht werden, doch korrelieren die Trends nicht mit der Amplifikation, da sie unabhängig sind von der Templatmenge und auch auftreten, wenn dem Reaktionsansatz die DNA-Polymerase, die Primer oder die dNTPs fehlen und es so zu keinerlei Amplifikation kommen kann. Werden diese Trends nicht korrigiert, korrelieren die Steigungen der log-linearen Abschnitte der Signalkurven mit der Stärke der Trends: Je größer der durch den Trend bedingte Signalanteil während der beobachteten log-linearen Phase, desto kürzer und flacher ist diese. Da sich die trendbedingte Signaländerung bei Kurven mit später exponentieller Phase (also geringerer Templatmenge) stärker auswirkt, flachen die log-linearen Abschnitte innerhalb einer Verdünnungsserie mit abnehmender Templatmenge ab (siehe **Abb. 3-14 B**, S. 67). Da die Steilheit dieser Phase ein direktes Maß für die Amplifikationseffizienz darstellt, impliziert diese Beobachtung, dass die Amplifikationseffizienz umgekehrt proportional zur Templatmenge ist, was biochemisch nicht begründbar ist. Im Gegenteil: man würde sogar erwarten, dass wenige Templatkopien sogar mit höherer Effizienz amplifiziert werden sollten, da der durchschnittliche Überschuss an Reaktionskomponenten größer ist und die Wahrscheinlichkeit einer Rehybridisierung der Template geringer ist. Die mit abnehmender Templatmenge flacher werdenden log-linearen Abschnitte der Signalkurven bedingen weiterhin, dass die anhand der Kalibriergeraden ermittelte Effizienz (siehe **Gl. 1.6**) mit größer werdenden Schwellenwerten abnimmt (**Abb. 3-18**, S. 72). Nach der in **3.2.1.2** beschriebenen Korrektur der Signaltrends sind die log-linearen Abschnitte der Signalkurven, unabhängig von der Templatmenge, parallel (siehe **Abb. 3-14 D**, S. 67) und die anhand der Kalibriergeraden ermittelte Effizienz ist nahezu unabhängig vom Schwellenwert. Zudem stimmen die anhand der Steigungen der log-linearen Abschnitte der individuellen Signalkurven berechneten Amplifikationseffizienzen sehr gut mit der aus der Kalibriergeraden ermittelten Effizienz überein (**Abb. 3-18**, S. 72). Dies ist bei nicht korrigierten Daten meistens nicht der Fall, was die Limitation der konventionellen Methoden nachdrücklich demonstriert. Diese Effekte belegen, dass die Korrektur um unspezifische Trends sinnvoll ist und die für Quantifizierungen wesentlichen Charakteristika der Signalkurven herausstellt.

Der entwickelte Algorithmus zur quantitativen Auswertung von RT-PCR-Daten erhöht die Reproduzierbarkeit, die Präzision sowie auch die Genauigkeit von Quantifizierungen beträchtlich, ohne den dynamischen Bereich einzuschränken (**3.2.1.6**). Dies ermöglicht den verlässlichen Einsatz der RT-PCR für die Detektion und Quantifizierung von Genamplifikationen und -deletionen, wie in **3.3** exemplarisch an den Bestimmungen der genomischen

Frequenz eines Retrotransposons, der Amplifikation des humanen Onkogens *HER-2/neu* und schließlich der Deletion des humanen Tumorsuppressorgens *p53* gezeigt wurde. Anhand einer neu entwickelten Methode zur absoluten Quantifizierung von *Single-copy-*Genen mittels der RT-PCR wurden außerdem die Genomgrößen verschiedener eukaryoter Modellorganismen bestimmt (**3.4.1**). Die ermittelte Genomgröße des sequenzierten Stamms der Bäckerhefe (*S. cerevisiae*, Stamm #368 FY 1679) von 12.1 MBp stimmt exakt mit dem aus der Sequenz abgeleiteten Wert überein, [247] was die hohe Genauigkeit dieses Verfahrens belegt. Die Auswertung derselben Messdaten mit dem Auswerteprogramm des LightCyclers (LCDA, Version 3.5) ergab für manche Experimente deutlich abweichende Ergebnisse. So wurde z. B. die Genomgröße des Menschen zu etwa 2.0 GBp bestimmt, was ungefähr 30 % unter den aktuellen, auf den Ergebnissen der Sequenzierprojekte basierenden Schätzungen von 2.8-3.0 GBp liegt, während die Auswertung mit den in dieser Arbeit entwickelten Algorithmen einen Wert von 2.9 GBp liefert.

4.5 Schmelzkurvenanalyse

Die in der PCR entstandenen PCR-Produkte können zudem im Anschluss an die Amplifikation mittels einer Schmelzkurvenanalyse charakterisiert werden, z. B. um bei SYBR Green I-Detektion zu bestimmen, ob spezifisches Produkt gebildet wurde [97] oder um Sequenzvarianten zu identifizieren bzw. Mutationen nachzuweisen.[98,99,102,252] Dazu wird die Fluoreszenz kontinuierlich aufgezeichnet, während die Temperatur der Reaktionsansätze langsam erhöht wird. Bei Überschreiten der Schmelztemperatur des detektierten Produkts fällt das Signal, durch den kooperativen Prozesses des Aufschmelzens von dsDNA, stark ab. Diejenige Temperatur, bei welcher die Signalabnahme maximal ist, wird definitionsgemäß als Schmelztemperatur bezeichnet. Die Menge geschmolzenen Produkts ist näherungsweise proportional zum Betrag des Signalunterschieds vor und nach dem Schmelzprozess. Da über die Schmelzkurven nicht nur PCR-Produkte mit verschiedenen Schmelztemperaturen, z. B. aufgrund unterschiedlicher Längen oder Sequenzen, innerhalb eines Reaktionsansatzes unterschieden werden können, sondern auch das Verhältnis ihrer Mengen bestimmt werden kann, lässt sich die Schmelzkurvenanalyse auch zur Quantifizierung von Templatmengen nach kompetitiver Amplifikation einsetzen.[147,217] Zudem kann das Mengenverhältnis zweier Gene in einer Probe einfach bestimmt werden, um beispielsweise Amplifikationen und Deletionen von Genen zu identifizieren.[253] Bei der Abschätzung der PCR-Produktmengen anhand der Schmelzkurven tritt jedoch ein systematischer Fehler auf, da die Änderungen des Fluoreszenzsignals nicht ausschließlich von den Schmelzprozessen hervorgerufen werden, sondern schon alleine die Erhöhung der Temperatur die Fluoreszenzausbeute deutlich verringert.[216] Die Stärke dieses als temperaturabhängiger Quench bezeichneten Prozesses ist seinerseits abhängig von der Gesamtfluoreszenz bei der jeweiligen Temperatur, also vom Anteil der noch nicht geschmolzenen Produkte. Aus diesem Grund werden in einem Gemisch die relativen Mengen desjenigen Produkts mit der höheren Schmelztemperatur unterbestimmt, und zwar umso stärker, je geringer sein Anteil an der Gesamtmenge ist (siehe **Abb. 3-23**, S. 79). In

Abschnitt **0** dieser Arbeit wird erstmals ein Verfahren vorgestellt, welches die Anteile des temperaturbedingten Quenchs identifiziert und die Schmelzkurven entsprechend korrigiert. Dadurch konnte sowohl der systematische Fehler eliminiert (siehe **Abb. 3-23**, S. 79) als auch die Sensitivität bzw. der dynamische Bereich der Analysen vergrößert werden, was die Eignung der kompetitiven Schmelzkurvenanalyse für relative Quantifizierungen erhöht. Weiterhin widerlegen diese Ergebnisse die Vermutungen von AL-ROBAIY *et al.*[217] und RIRIE *et al.*,[97] dass die Abweichungen der Mengenbestimmungen bei der Schmelzkurvenanalyse zu den realen Werten bedingt sei durch eine unterschiedliche spezifische Fluoreszenz von SYBR Green I in Abhängigkeit von der Länge und der Sequenz des jeweiligen PCR-Produkts.

4.6 Software zur Auswertung von Real-time-PCR-Daten

Mit der zunehmend häufigeren Anwendung der RT-PCR in der klinischen Routinediagnostik ist es von nicht zu unterschätzender Bedeutung, dass die Messdaten korrekt, zuverlässig und nach Möglichkeit vollständig automatisch ausgewertet werden, um eine willkürliche und oft inadäquate Beeinflussung der Ergebnisse zu minimieren. Hierzu sind entsprechende automatisierbare Auswerteverfahren notwendig, in Form einer leicht bedienbaren, anwenderfreundlichen Software. Die in dieser Arbeit entwickelte Software erfüllt diese Kriterien; sie kann ohne weitere EDV-Kenntnisse bedient werden und liefert vollautomatisch die für die jeweiligen Daten optimalen Ergebnisse (**3.2.4**).

Durch die Art der Auswertung wird sichergestellt, dass in reproduzierbarer Weise immer die bestmöglichen Resultate erzielt werden. Da die Ergebnisse nicht mehr von willkürlichen Einstellungen der Experimentatoren abhängen, sind die Messergebnisse gleicher Experimente aus verschiedenen Arbeitsgruppen untereinander besser vergleichbar, was die Validierung von Analyseergebnissen vereinfacht. Ein weiterer entscheidender Vorteil dieser Software gegenüber den Programmen der Gerätehersteller ergibt sich aus der automatischen Detektion der exponentiellen Phasen von Amplifikationskurven, wodurch auch kritische Signalverläufe erkannt werden können und auswertbar werden. Das Programm ist somit in der Lage, den Benutzer entsprechend zu informieren bzw. zu warnen und Signalkurven, die keine erkennbare exponentielle Phase aufweisen, von der quantitativen Auswertung auszuschließen. Schließlich wird noch eine Bewertung der Vertrauenswürdigkeit der Ergebnisse vorgenommen, indem die Charakteristika der exponentiellen Phasen der Reaktionsansätze, insbesondere die daraus ermittelten Effizienzen für die individuellen Ansätze, untereinander verglichen werden. Dies erhöht die Tauglichkeit für den Routineeinsatz sowie bei der Auswertung von großen Datenmengen, wie sie beispielsweise bei Probenhochdurchsatzstudien anfallen, beträchtlich.

Für spezielle Routineanwendungen sollten zukünftige Programme modular aufgebaut sein, damit sie möglichst einfach und gut an die speziellen Anforderungen der jeweiligen Experimente und Auswertungen angepasst werden können. Damit können die Programme direkt bei der Planung, Durchführung und Auswertung von Experimenten assistieren und so

mögliche Fehler, wie sie z. B. durch eine nicht korrekte Interpretation der Messdaten auftreten, vermeiden helfen.

4.7 Das methodische Potenzial der Real-time-PCR

Das Prinzip der RT-PCR birgt ein großes methodisches Potenzial, insbesondere in der Kombination mit anderen Techniken wie der kompetitiven PCR oder der allelspezifischen PCR.[98,101,123,147,217,253,255,257] Die kompetitive PCR gilt derzeit als das genaueste PCR-basierte Verfahren zur Quantifizierung von Nukleinsäuren mit Variationskoeffizienten von etwa 10 %,[205] jedoch wird ein brauchbarer Kompetitor benötigt und jede Probe muss in einer eigenen Serie von Einzelreaktionen mit unterschiedlichen Konzentrationen des Kompetitors untersucht werden, um den Äquivalenzpunkt von Kompetitor und Probe zu bestimmen.[17,19,20,256] Die analytische Unterscheidung und Mengenbestimmung der PCR-Produkte mit Kompetitor- bzw. Probensequenz ist bei konventionellen, zumeist auf elektrophoretischen Trennungen und auf Densitometrie basierenden Verfahren aufwändig und fehleranfällig. In Kombination mit der RT-PCR können die relativen Mengen der PCR-Produkte von Kompetitor und Probe, deren Sequenzen sich nur minimal zu unterscheiden brauchen, sehr einfach und schnell durch Schmelzkurvenanalysen bestimmt werden.[147,217,253] Allerdings ist der dynamische Bereich nur etwa 1:10 (Mengenverhältnis von Kompetitor zu Probe bzw. deren inverses Verhältnis, das noch mit einem akzeptablen Fehler bestimmt werden kann) und die Genauigkeit normalerweise nicht besser als bei Quantifizierungen mittels externer Standards,[217] da die Schmelzkurvendaten verrauscht sind und der temperaturabhängige Quench die schmelzprozessbedingten Fluoreszenzänderungen maskiert (siehe **3.2.2.2**). Wenn statt der Schmelzkurven wie in **3.4.2.4** gezeigt die Amplifikationskurven nach asymmetrischer PCR und sequenzspezifischer Detektion bei stringenter und nicht-stringenter Temperatur analysiert werden, kann die Genauigkeit deutlich erhöht werden, so dass Variationskoeffizienten von weniger als 3 % erreicht werden (**3.4.2.6**). Dies ist eine Genauigkeit in der Bestimmung von Nukleinsäuren, die derzeit von keinem anderen PCR-basierten Verfahren erreicht wird. Möglich ist dies durch die Auswertung der Signalhöhen bzw. der Plateauwerte anstelle der C_T-Werte. Zum einen sind die Fehler der (relativen) Konzentrationsbestimmung hier nur linear abhängig vom Fehler der Signalhöhenbestimmung, zum anderen lassen sich viele Datenpunkte einer Amplifikationskurve mitteln, um den durch das Signalrauschen bedingten Messfehler zu verkleinern, während bei der Konzentrationsbestimmung über die C_T-Werte dazu mehrere Reaktionen parallel durchgeführt werden müssen, um entsprechend viele C_T-Werte ermitteln zu können. Weiterhin machen die Ergebnisse klar die Vor- und Nachteile der Verwendung der C_T-Werte zur Quantifizierung deutlich. Ein hoher dynamischer Bereich erlaubt nur eine vergleichsweise geringe Genauigkeit. Bei Verwendung der Signalintensitäten statt der C_T-Werte hingegen ist der dynamische Bereich wieder eingeschränkt, die Genauigkeit jedoch sehr hoch. So gibt es nicht die generell beste Methode zur Quantifizierung, sondern die Wahl der Methode ergibt sich aus der jeweiligen Problemstellung. Ein großer dynamischer Bereich bei gleichzeitig hoher Genauigkeit kann im Bedarfsfall durch eine Kombination entsprechender Methoden erreicht werden.

Aufgrund des potenziell sehr großen dynamischen Bereichs der RT-PCR für quantitative Bestimmungen eignet sich das Verfahren prinzipiell für den Einsatz in der Analyse von Mischproben, die eine große Varianz der Mischungsverhältnisse aufweisen. Dies ist sinnvoll z. B. bei Chimärenanalysen,[123] populationsgenetischen Studien[257] und in der Lebensmittelanalytik, um beispielsweise die Sortenzusammensetzung verarbeiteter pflanzlicher Nahrungsmittel kontrollieren zu können. Ein besonderes Problem stellt hier der hohe Grad an Ähnlichkeit der Sequenzen in den Mischproben dar. Verfahren wie STR-, VNTR- und RFLP-Analysen, welche die Sequenzvarianten gemeinsam nachweisen bzw. analysieren, können Konzentrationsunterschiede von mehr als etwa 1:10 nicht mehr sicher auflösen,[126] ihr dynamischer Bereich ist also recht klein. Höhere Dynamik kann durch allel-spezifische Amplifikationen erreicht werden, die jedoch gerade bei nur geringen Sequenzunterschieden häufig zu falschen oder nicht reproduzierbaren Ergebnissen führen.[257] In der Kombination mit der RT-PCR sind die quantitativen Aussagen verlässlich, ohne den dynamischen Bereich einzuschränken. Die in dieser Arbeit (**3.4.3**) mittels allelspezifischer RT-PCR für zwei exemplarische Polymorphismen mit klinischer Relevanz bestimmten dynamischen Bereiche liegen um ein Vielfaches über denen von üblicherweise für derartige Fragestellungen verwendeten STR-Analysen, und das mit vergleichbarer Reproduzierbarkeit: Für Polymorphismen, die mittels allelspezifischer PCR im Allgemeinen extrem schwierig zu unterscheiden sind, ist der dynamische Bereich der hier entwickelten Methode mindestens um einen Faktor drei größer als bei den konventionell verwendeten Verfahren, bei einfach zu unterscheidenden Polymorphismen sogar mehr als 150mal größer.

4.8 Die Zukunft der Real-time-PCR

Die in dieser Arbeit vorgestellten Anwendungsbeispiele (**3.3**) und die methodischen Entwicklungen (**3.4**) vermitteln einen Eindruck über die zukünftigen Einsatzgebiete und Möglichkeiten der RT-PCR. Die schnelle, einfache und sichere Quantifizierung von Nukleinsäuren macht die RT-PCR zum Verfahren der Wahl in der klinischen Diagnostik von Infektionen und Tumoren,[272,273] wo man zukünftig besser in der Lage sein wird, den Erfolg von Therapien und Operationen zu verfolgen und frühzeitig Informationen über die adäquate Behandlung zu erhalten. Im Prinzip wird es möglich sein, noch während einer Operation wichtige Diagnosen zu stellen, z. B. über die genaue Art und das Stadium eines Tumors, so dass gegebenenfalls die Operations- und Behandlungsstrategie angepasst werden kann, vergleichbar dem Schnellschnitt für eine histopathologische Untersuchung.

Die RT-PCR wird verstärkt helfen, die genetischen Prozesse bei der Zelldifferenzierung,[297] bei der Tumorgenese[298] und bei Immunreaktionen[280,299] zu entschlüsseln und zu verstehen. Da bei Differenzierungsprozessen während der Ontogenese sowie der Tumorentstehung der meisten Eukaryonten, die Genregulation durch den Methylierungsgrad der DNA eine entscheidende Rolle spielt, werden verstärkt Methyl-DNA-sensitive Varianten der quantitativen RT-PCR entwickelt werden.[298,300-303] Die Aktivität von Genen, die in Expressionsstudien anhand von Transkriptionsprofilen aus *Microarray*-Analysen identifiziert wurden, lässt sich durch den Einsatz von RT-PCR-basierten Verfahren kostengünstig und schnell in Abhängig-

keit verschiedener Parameter bestimmen und verifizieren.[288-290] Die RT-PCR ist daher heute die Referenzmethode zur Bestätigung von *Microarray*-Daten. Weiterhin werden quantitative SNP-Analysen mit großem dynamischen Messbereich vereinfacht, was die Anwendung der RT-PCR bei der Mutationsdetektion,[43,75,94,252] bei der Chimärenanalytik[257] sowie auch bei popolationsgenetischen Fragestellungen attraktiv macht.[121,123,304]

Wesentliche Beiträge zu zukünftigen biologischen Forschungsfragen wird die RT-PCR auch durch ihre Kombination mit anderen fortgeschrittenen Techniken wie der bereits erwähnten *Microarray*-Analyse zur Erstellung von Transkriptionsprofilen,[274,288-290,305] der Mikrodissektion[306-308] oder der DNA-Extraktion aus paraffinfixierten Proben leisten.[146,148,309,310] So wird es zunehmend einfacher, die Genexpression oder die genomischen Genkopienzahlen in Proben zu bestimmen, die nur sehr wenige Zellen oder nur bestimmte Zelltypen enthalten, was neue Einblicke in die Differenzierungsprozesse von Tumoren und Geweben geben kann. In dieser Arbeit sind Anstöße gegeben, wie die RT-PCR selbst verbessert und für neue Fragestellungen mit höheren Anforderungen an Präzision, Genauigkeit und Dynamik angepasst und eingesetzt werden kann. Die dazu vorgestellten Beispiele können direkt in der DNA-Analytik eingesetzt werden und bereichern auf Grund ihrer medizinischen Relevanz unmittelbar die medizinische Forschung und klinische Labordiagnostik. Die in dieser Arbeit entwickelte Software schafft weiterhin allgemein in diesem Gebiet die Grundlage für eine erhebliche Steigerung der Leistungsfähigkeit, Verlässlichkeit und Genauigkeit der Ergebnisse von RT-PCR-Experimenten.

5 ZUSAMMENFASSUNG

Derjenige ist wahnsinnig, der ein Geheimnis in jedweder anderer Form niederschreibt als in einer, die es vor den Gewöhnlichen verbirgt und selbst den Gelehrten und ernsthaften Studenten nur mit Mühe sinnhaftig werden lässt.

Robert Bacon „Über die Nichtigkeit der Magie"

Die molekulare DNA-Analytik ist von großer Bedeutung für die biologische und medizinische Forschung sowie zunehmend auch für die klinische Diagnostik. Hier bieten real-time-PCR-basierte Verfahren neue Möglichkeiten der schnellen und verlässlichen Quantifizierung genetischer Elemente und der einfachen Mutationsdetektion. Die vorliegende Arbeit behandelt einen Querschnitt durch das Spektrum der real-time-PCR-basiertern DNA-Analytik und gibt einen umfassenden Überblick über die theoretischen und biochemischen Grundlagen, zeigt die Einsatzmöglichkeiten für DNA-analytische Fragestellungen und stellt neue methodische Entwicklungen vor, die neue Standards bzw. Grenzen für Genauigkeit und Dynamik setzen.

Die systematische Untersuchung der Einflüsse biochemischer und technischer Parameter auf das Ergebnis von Messungen in Kombination mit Analysen theoretischer Modelle der Real-time-PCR-Prozesse erlaubt erstmals die Bestimmung möglicher Schwachstellen sowie der Grenzen für Präzision und Genauigkeit von Analysesystemen. So konnte gezeigt werden, dass die theoretisch erreichbare Genauigkeit von Quantifizierungen über C_T-Werte des hier verwendeten LightCyclers bei etwa 10 % liegt. Durch die Verwendung neuer, in dieser Arbeit entwickelter Algorithmen zur Datenauswertung wird diese theoretische Grenze nahezu erreicht: Die Fehler können von mehr als 50 % nach konventioneller Auswertung auf deutlich unter 15 % gesenkt werden. Die Anwendung dieser Verfahren wurde an drei ausgewählten Fragestellungen mit tumordiagnostischem Fokus demonstriert (XIR-2, *HER-2/neu* und *p53*), die sehr unterschiedliche Ansprüche an den dynamischen Bereich und die Präzision der Quantifizierung stellen. Dabei übertrifft insbesondere die Genauigkeit des Verfahrens zur Bestimmung des Deletionsgrades von *p53*-Genkopien jene von anderen analytischen Methoden deutlich. Bereits der Verlust eines der beiden Allele in lediglich 25 % der Zellen einer Probe kann hochsignifikant nachgewiesen werden.

Die Algorithmen zur quantitativen Auswertung von Real-time-PCR-Daten erlauben erstmals eine optimale und vollautomatische Auswertung. Durch ein hier erstmals beschriebenes Verfahren zur Korrektur temperaturabhängiger Signaländerungen konnten die Sensitivität und die Genauigkeit von Schmelzkurvenanalysen beträchtlich erhöht werden. Die beschriebenen Algorithmen wurden in einer neuen, sehr anwenderfreundlichen Software (SoFAR) implementiert.

Schließlich werden mehrere neuartige Verfahren zum Einsatz der Real-time-PCR vorgestellt: Ein Verfahren erlaubt die sehr genaue Bestimmung von Genomgrößen, wie am Beispiel von *S. cerevisiae* (12.1 MBp), *H. sapiens* (2.9 GBp) und *X. maculatus* (551 MBp) gezeigt ist. Ein weiteres Verfahren beruht auf einer allelspezifischen Real-time-PCR und kann zur

einfachen Genotypisierung sowie zur Chimärenanalytik eingesetzt werden, wie am Beispiel der klinisch relevanten SNPs von *PAI-1* und *F7c* gezeigt. Die erreichten dynamischen Bereiche zur Unterscheidung kleiner Polymorphismen sind hier über 100mal größer als die konventioneller Verfahren. Mit dem neuen Verfahren zur kompetitiven Real-time-PCR sind hochgenaue Konzentrationsbestimmungen möglich. Mit dieser Methode wurde die Menge der cDNA von *BCR/ABL*-Fusionstranskripten in Einzelbestimmungen mit einem Fehler von weniger als 3 % bestimmt.

Die Ergebnisse dieser Arbeit steigern die Leistungsfähigkeit, Verlässlichkeit und Genauigkeit der Ergebnisse von Real-time-PCR-Experimenten und erweitern die möglichen Anwendungsbereiche der Real-time-PCR für DNA-analytische Fragestellungen. Die entwickelten Verfahren bereichern das methodische Arsenal der molekularen Genetik und Medizin.

6 LITERATURVERZEICHNIS

1. MORRISON, C. & GANNON, F. (1994) The impact of the PCR plateau phase on quantitative PCR. *Biochim Biophys Acta* **1219**: 493-498.

2. KLEPPE, K., OHTSUKA, E., KLEPPE, R., MOLINEUX, I. & KHORANA, H.G. (1971) Studies on Polynukleotides. XCVI. Repair of short synthetic DNA's as catalysed by DNA polymerases. *J Mol Biol* **56**: 341-361.

3. SAIKI, R.K., GELFAND, D.H., STOFFEL, S., SCHARF, S.J., HIGUCHI, R., *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.

4. WONG, C., DOWLING, C.E., SAIKI, R.K., HIGUCHI, R.G., ERLICH, H.A., *et al.* (1987) Characterization of beta-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. *Nature* **330**: 384-386.

5. SAIKI, R.K., SCHARF, S., FALOONA, F., MULLIS, K.B., HORN, G.T., *et al.* (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.

6. ABBOTT, M.A., POIESZ, B.J., BYRNE, B.C., KWOK, S., SNINSKY, J.J., *et al.* (1988) Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified in vitro. *J Infect Dis* **158**: 1158-1169.

7. ERLICH, H.A. (1989) Polymerase chain reaction. *J Clin Immunol* **9**: 437-447.

8. BLOCH, W. (1991) A biochemical perspective of the polymerase chain reaction. *Biochem* **30**: 2735-2247.

9. LINZ, U. & DEGENHARDT, H. (1990) Die Polymerase Kettenreaktion. *Naturwissenschaften* **77**: 515-530.

10. ARNHEIM, N. & ERLICH, H. (1992) Polymerase chain reaction strategy. *Annu Rev Biochem* **61**: 131-156.

11. BEJ, A.K., MAHBUBANI, M.H. & ATLAS, R.M. (1991) Amplification of nucleic acids by polymerase chain reaction (PCR) and other methods and their applications. *Crit Rev Biochem Mol Biol* **26**: 301-334

12. EHRLICH, H.A. (1989) *PCR Technology. Principles and applications for DNA amplification.*, Stockton Press, New York, London.

13. INNIS, M.A., GELFAND, D.H. & SININSKY, J.J. (1995) *PCR Strategies.* Academic Press, San Diego, London.

14. MULLIS, K.B., FERRÉ, F. & GIBBS, R.A. (1994) *The polymerase chain reaction.* Birkhäuser, Basel, Berlin, Boston.

15. VILLARREAL, X.C., GRANT, B.W. & LONG, G.L. (1991) Demonstration of osteonectin mRNA in megakaryocytes: the use of the polymerase chain reaction. *Blood* **78**: 1216-1222.

16. SYKES, P.J., NEOH, S.H., BRISCO, M.J., HUGHES, E., CONDON, J., *et al.* (1992) Quantitation of targets for PCR by use of limiting dilution. *Biotechniques* **13**: 444-449.

17. GILLILAND, G., PERRIN, S., BLANCHARD, K. & BUNN, H.F. (1990) Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction. *Proc Natl Acad Sci USA* **87**: 2725-2729.

18. GOERKE, C., BAYER, M.G. & WOLZ, C. (2001) Quantification of bacterial transcripts during infection using competitive reverse transcription-PCR (RT-PCR) and LightCycler RT-PCR. *Clin Diagn Lab Immunol* **8**: 279-282.

19. RAEYMAEKERS, L. (1995) A commentary on the practical applications of competitive PCR. *Genome Res* **5**: 91-94.

20. ZIMMERMANN, K. & MANNHALTER, J.W. (1996) Technical aspects of quantitative competitive PCR. *Biotechniques* **21**: 268-272, 274-269.

21. O'CONNELL, J., GOODE, T. & SHANAHAN, F. (1998) Quantitative measurement of mRNA expression by competitive RT-PCR. *Methods Mol Biol* **92**: 183-193.

22. HIGUCHI, R., FOCKLER, C., DOLLINGER, G. & WATSON, R. (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (NY)* **11**: 1026-1030.

23. HEID, C.A., STEVENS, J., LIVAK, K.J. & WILLIAMS, P.M. (1996) Real time quantitative PCR. *Genome Res* **6**: 986-994.

24. NAZARENKO, I.A., BHATNAGAR, S.K. & HOHMAN, R.J. (1997) A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Res* **25**: 2516-2521.

25. LE-PECQ, J.-B. & PAOLETTI, C. (1966) A new fluorimetric method for RNA and DNA determination. *Anal Biochem* **17**: 100-107.

26. WITTWER, C.T., HERRMANN, M.G., MOSS, A.A. & RASMUSSEN, R.P. (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques* **22**: 130-131, 134-138.

27. MORRISON, T.B., WEIS, J.J. & WITTWER, C.T. (1998) Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques* **24**: 954-958, 960, 962.

28. ISHIGURO, T., SAITOH, J., YAWATA, H., YAMAGISHI, H., IWASAKI, S., *et al.* (1995) Homogeneous quantitative assay of hepatitis C virus RNA by polymerase chain reaction in the presence of a fluorescent intercalater. *Anal Biochem* **229**: 207-213.

29. TSENG, S.Y., MACOOL, D., ELLIOT, V., TICE, G., JACKSON, R., *et al.* (1997) A homogeneous fluorescence polymerase chain reaction assay to identify Salmonella. *Anal Biochem* **245**: 207-212.

30. JIN, X., YUE, S., WELLS, K.S. & SINGER, V.L. (1994) SYBR Green I: a new fluorescent dye optimized for detection of picogram amounts of DNA in gels. *Biophys J* **66**: A159.

31. SINGER, V.L., JIN, X., RYAN, D. & YUE, S. (1994) SYBR Green dyes: ultrasensitive stains for detection of DNA and RNA in electrophoretic gels. *Biomed Products* **19**: 86-72.

32. CARDULLO, R.A., AGRAWAL, S., FLORES, C., ZAMECNIK, P.C. & WOLF, D.E. (1988) Detection of nucleic acid hybridization by nonradiative fluorescence resonance energy transfer. *Proc Natl Acad Sci USA* **85**: 8790-8794.

33. WU, P. & BRAND, L. (1994) Resonance energy transfer: Methods and applications. *Anal Biochem* **218**: 1-13.

34. CLEGG, R.M. (1995) Fluorescence resonance energy transfer. *Curr Opin Biotechnol* **6**: 103-110.

35. FÖRSTER, T. (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann Phys* **6**: 55-75.

36. STRYER, L. & HAUGLAND, R.P. (1967) Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci USA* **58**: 719-726.

37. CLEGG, R.M. (1992) Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol* **211**: 353-388.

38. HOLLAND, P.M., ABRAMSON, R.D., WATSON, R. & GELFAND, D.H. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'→3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci USA* **88**: 7276-7280.

39. GIBSON, U.E., HEID, C.A. & WILLIAMS, P.M. (1996) A novel method for real time quantitative RT-PCR. *Genome Res* **6**: 995-1001.

40. KOSTRIKIS, L.G., TYAGI, S., MHLANGA, M.M., HO, D.D. & KRAMER, F.R. (1998) Spectral genotyping of human alleles. *Science* **279**: 1228-1229.

41. TYAGI, S. & KRAMER, F.R. (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* **14**: 303-308.

42. TYAGI, S., BRATU, D.P. & KRAMER, F.R. (1998) Multicolor molecular beacons for allele discrimination. *Nat Biotechnol* **16**: 49-53.

43. WHITCOMBE, D., BROWNIE, J., GILLARD, H.L., MCKECHNIE, D., THEAKER, J., *et al.* (1998) A homogeneous fluorescence assay for PCR amplicons: its application to real-time, single-tube genotyping. *Clin Chem* **44**: 918-923.

44. WHITCOMBE, D., THEAKER, J., GUY, S.P., BROWN, T. & LITTLE, S. (1999) Detection of PCR products using self-probing amplicons and fluorescence. *Nat Biotechnol* **17**: 804-807.

45. ISACSON, J., CAO, H., OHLSSON, L., NORDGREN, S., SVANVIC, N., *et al.* (2000) Rapid and specific detection of PCR products using light-up probes. *Mol Cell Probes* **14**: 321-328.

46. SVANVIC, N., SEHLSTEDT, U., SJÖBACK, R. & KUBISTA, M. (2000) Detection of PCR products in real time using light-up probes. *Anal Biochem* **287**: 179-182.

47. SVANVIC, N., WESTMAN, G., WANG, D. & KUBISTA, M. (2001) Light-up probes: thiazole orange-conjugated peptide nucleic acid for the detection of target nucleic acid in homogeneous solution. *Anal Biochem* **281**: 26-35.

48. KURATA, S., KANAGAWA, T., YAMADA, K., TORIMURA, M., YOKOMAKU, T., *et al.* (2001) Fluorescent quenching-based quantitative detection of specific DNA/RNA using a BODIPY FL-labelled probe or primer. *Nucleic Acids Res* **29**: e34.

49. CROCKETT, A.O. & WITTWER, C.T. (2001) Fluorescein-labelled oligonucleotides for real-time PCR: using inherent quenching of deoxyguanosine nucleotides. *Anal Biochem* **281**: 179-182.

50. WHITCOMB, D.C., GORRY, M.C., PRESTON, R.A., FUREY, W., SOSSENHEIMER, M.J., *et al.* (1996) Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet* **14**: 141-145.

51. YAMADA, N., ARINAMI, T., YAMAKAWA-KOBAYASHI, K., WATANABE, H., SOHDA, S., *et al.* (2000) The 4G/5G polymorphism of the plasminogen activator inhibitor-1 gene is associated with severe preeclampsia. *J Hum Genet* **45**: 138-141.

52. COTTON, E.A., ALLSOP, R.F., GUEST, J.L., FRAZIER, R.R., KOUMI, P., *et al.* (2000) Validation of the AMPFISTR SGM plus system for use in forensic casework. *Forensic Sci Int* **112**: 151-161.

53. HOHOFF, C. & BRINKMANN, B. (1999) Human identity testing with PCR-based systems. *Mol Biotechnol* **13**: 123-136.

54. LITT, M. & LUTY, J.A. (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* **44**: 397-401.

55. LEFRERE, J.J., MARIOTTI, M., WATTEL, E., LEFRERE, F., INCHAUSPE, G., *et al.* (1992) Towards a new predictor of AIDS progression through the quantitation of HIV-1 DNA copies by PCR in HIV-infected individuals. *Br J Haematol* **82**: 467-471.

56. KRINGS, M., STONE, A., SCHMITZ, R.W., KRAINITZKI, H., STONEKING, M., *et al.* (1997) Neandertal DNA sequences and the origin of modern humans. *Cell* **90**: 19-30.

57. TAUTZ, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* **17**: 6463-6471.

58. VENTER, J.C., ADAMS, M.D., MYERS, E.W., LI, P.W., MURAL, R.J., *et al.* (2001) The sequence of the human genome. *Science* **291**: 1304-1351.

59. SACHIDANANDAM, R., WEISSMAN, D., SCHMIDT, S.C., KAKOL, J.M., STEIN, L.D., *et al.* (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**: 928-933.

60. MULLIKIN, J.C., HUNT, S.E., COLE, C.G., MORTIMORE, B.J., RICE, C.M., *et al.* (2000) An SNP map of human chromosome 22. *Nature* **407**: 516-520.

61. LI, W.H. & SADLER, L.A. (1991) Low nucleotide diversity in man. *Genetics* **129**: 513-523.

62. KAN, Y.W. & DOZY, A.M. (1978) Antenatal diagnosis of sickle-cell anaemia by D.N.A. analysis of amniotic-fluid cells. *Lancet* **2**: 910-912.

63. SANGER, F., NICKLEN, S. & COULSON, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463-5467.

64. SYVANEN, A.C., AALTO-SETALA, K., HARJU, L., KONTULA, K. & SODERLUND, H. (1990) A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **8**: 684-692.

65. BOLAND, E.J., PILLAI, A., ODOM, M.W. & JAGADEESWARAN, P. (1994) Automation of the Maxam-Gilbert chemical sequencing reactions. *Biotechniques* **16**: 1088-1092, 1094-1085.

66. DOLAN, M., ALLY, A., PURZYCKI, M.S., GILBERT, W. & GILLEVET, P.M. (1995) Large-scale genomic sequencing: optimization of genomic chemical sequencing reactions. *Biotechniques* **19**: 264-268, 270-264.

67. TULLY, G., SULLIVAN, K.M., NIXON, P., STONES, R.E. & GILL, P. (1996) Rapid detection of mitochondrial sequence polymorphisms using multiplex solid-phase fluorescent minisequencing. *Genomics* **34**: 107-113.

68. GLAVAC, D. & DEAN, M. (1995) Applications of heteroduplex analysis for mutation detection in disease genes. *Hum Mutat* **6**: 281-287.

69. FISCHER, S.G. & LERMAN, L.S. (1983) DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci USA* **80**: 1579-1583.

70. FODDE, R. & LOSEKOOT, M. (1994) Mutation detection by denaturing gradient gel electrophoresis (DGGE). *Hum Mutat* **3**: 83-94.

71. RIESNER, D., STEGER, G., WIESE, U., WULFERT, M., HEIBEY, M., *et al.* (1992) Temperature gradient gel electrophoresis for the detection of polymorphic DNA and quantitative polymerase chain reaction. *Electrophoresis* **13**: 632-636.

72. HENCO, K., HARDERS, J., WIESE, U. & RIESNER, D. (1994) Temperature gradient gel electrophoresis (TGGE) for the detection of polymorphic DNA and RNA. *Methods Mol Biol* **31**: 211-228.

73. UNDERHILL, P.A., JIN, L., LIN, A.A., MEHDI, S.Q., JENKINS, T., *et al.* (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* **7**: 996-1005.

74. O'DONOVAN, M.C., OEFNER, P.J., ROBERTS, S.C., AUSTIN, J., HOOGENDOORN, B., *et al.* (1998) Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. *Genomics* **52**: 44-49.

75. HECKER, K.H., TAYLOR, P.D. & GJERDE, D.T. (1999) Mutation detection by denaturing DNA chromatography using fluorescently labeled polymerase chain reaction products. *Anal Biochem* **272**: 156-164.

76. BABON, J.J., YOUIL, R. & COTTON, R.G. (1995) Improved strategy for mutation detection - a modification to the enzyme mismatch cleavage method. *Nucleic Acids Res* **23**: 5082-5084.

77. DEL TITO, B.J., JR., POFF, H.E., 3RD, NOVOTNY, M.A., CARTLEDGE, D.M., WALKER, R.I., 2ND, *et al.* (1998) Automated fluorescent analysis procedure for enzymatic mutation detection. *Clin Chem* **44**: 731-739.

78. MASHAL, R.D., KOONTZ, J. & SKLAR, J. (1995) Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nat Genet* **9**: 177-183.

79. PARSONS, B.L. & HEFLICH, R.H. (1997) Evaluation of MutS as a tool for direct measurement of point mutations in genomic DNA. *Mutat Res* **374**: 277-285.

80. SACHADYN, P., STANISAWSKA, A. & KUR, J. (2000) One tube mutation detection using sensitive fluorescent dyeing of MutS protected DNA. *Nucleic Acids Res* **28**: e36.

81. HIGUCHI, R., VON BEROLDINGEN, C.H., SENSABAUGH, G.F. & ERLICH, H.A. (1988) DNA typing from single hairs. *Nature* **332**: 543-546.

82. TODD, A.V., IRELAND, C.M., RADLOFF, T.J., KRONENBERG, H. & ILAND, H.J. (1991) Analysis of *N-ras* gene mutations in acute myeloid leukemia by allele specific restriction analysis. *Am J Hematol* **38**: 207-213.

83. KRISTIANSEN, B.E., FERMER, C., JENKINS, A., ASK, E., SWEDBERG, G., *et al.* (1995) PCR amplicon restriction endonuclease analysis of the chromosomal *dhps* gene of *Neisseria meningitidis*: a method for studying spread of the disease-causing strain in contacts of patients with meningococcal disease. *J Clin Microbiol* **33**: 1174-1179.

84. EGGERDING, F.A. (1995) A one-step coupled amplification and oligonucleotide ligation procedure for multiplex genetic typing. *PCR Methods Appl* **4**: 337-345.

85. FISCHER, G.F., FAE, I., MOSER, S., PETRASEK, M., BASSLER, H.M., *et al.* (1996) Ligation based HLA-B*27 typing. *Tissue Antigens* **48**: 148-152.

86. DELAHUNTY, C., ANKENER, W., DENG, Q., ENG, J. & NICKERSON, D.A. (1996) Testing the feasibility of DNA typing for human identification by PCR and an oligonucleotide ligation assay. *Am J Hum Genet* **58**: 1239-1246.

87. STONE, G.G., OBERST, R.D., HAYS, M.P., MCVEY, S. & CHENGAPPA, M.M. (1995) Combined PCR-oligonucleotide ligation assay for rapid detection of *Salmonella serovars*. *J Clin Microbiol* **33**: 2888-2893.

88. BARANY, F. (1991) The ligase chain reaction in a PCR world. *PCR Methods Appl* **1**: 5-16.

89. BARANY, F. (1991) Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc Natl Acad Sci USA* **88**: 189-193.

90. WIEDMANN, M., WILSON, W.J., CZAJKA, J., LUO, J., BARANY, F., *et al.* (1994) Ligase chain reaction (LCR)-overview and applications. *PCR Methods Appl* **3**: S51-64.

91. CHEN, X., LIVAK, K.J. & KWOK, P.Y. (1998) A homogeneous, ligase-mediated DNA diagnostic test. *Genome Res* **8**: 549-556.

92. NEWTON, C.R., GRAHAM, A., HEPTINSTALL, L.E., POWELL, S.J., SUMMERS, C., *et al.* (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* **17**: 2503-2516.

93. OKAYAMA, H., CURIEL, D.T., BRANTLY, M.L., HOLMES, M.D. & CRYSTAL, R.G. (1989) Rapid, nonradioactive detection of mutations in the human genome by allele-specific amplification. *J Lab Clin Med* **114**: 105-113.

94. KYGER, E.M., KREVOLIN, M.D. & POWELL, M.J. (1998) Detection of the hereditary hemochromatosis gene mutation by real-time fluorescence polymerase chain reaction and peptide nucleic acid clamping. *Anal Biochem* **260**: 142-148.

95. ORUM, H., NIELSEN, P.E., EGHOLM, M., BERG, R.H., BUCHARDT, O., *et al.* (1993) Single base pair mutation analysis by PNA directed PCR clamping. *Nucleic Acids Res* **21**: 5332-5336.

96. LIVAK, K.J. (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* **14**: 143-149.

97. RIRIE, K.M., RASMUSSEN, R.P. & WITTWER, C.T. (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal Biochem* **245**: 154-160.

98. AOSHIMA, T., SEKIDO, Y., MIYAZAKI, T., KAJITA, M., MIMURA, S., *et al.* (2000) Rapid detection of deletion mutations in inherited metabolic diseases by melting curve analysis with LightCycler. *Clin Chem* **46**: 119-122.

99. LIPSKY, R.H., MAZZANTI, C.M., RUDOLPH, J.G., XU, K., VYAS, G., *et al.* (2001) DNA melting analysis for detection of single nucleotide polymorphisms. *Clin Chem* **47**: 635-644.

100. GIESENDORF, B.A., VET, J.A., TYAGI, S., MENSINK, E.J., TRIJBELS, F.J., *et al.* (1998) Molecular beacons: a new approach for semiautomated mutation analysis. *Clin Chem* **44**: 482-486.

101. NAUCK, M., WIELAND, H. & MARZ, W. (1999) Rapid, homogeneous genotyping of the 4G/5G polymorphism in the promoter region of the *PAII* gene by fluorescence resonance energy transfer and probe melting curves. *Clin Chem* **45**: 1141-1147.

102. KRETZSCHMAR, E., VON AHSEN, N. & TROBISCH, H. (2001) Rapid genotyping of human platelet antigen 5 with fluorophore-labelled hybridization probes on the LightCycler. *Br J Haematol* **114**: 397-399.

103. LAREU, M., PUENTE, J., SOBRINO, B., QUINTANS, B., BRION, M., *et al.* (2001) The use of the LightCycler for the detection of Y chromosome SNPs. *Forensic Sci Int* **118**: 163-168.

104. WITTWER, C.T., HERRMANN, M.G., GUNDRY, C.N. & ELENITOBA-JOHNSON, K.S. (2001) Real-time multiplex PCR assays. *Methods* **25**: 430-442.

105. LAGAAIJ, E.L., CRAMER-KNIJNENBURG, G.F., VAN KEMENADE, F.J., VAN ES, L.A., BRUIJN, J.A., *et al.* (2001) Endothelial cell chimerism after renal transplantation and vascular rejection. *Lancet* **357**: 33-37.

106. DAVID, B., BERNARD, D., NAVENOT, J.M., MULLER, J.Y. & BLANCHARD, D. (1999) Flow cytometric monitoring of red blood cell chimerism after bone marrow transplantation. *Transfus Med* **9**: 209-217.

107. GAIGER, A., MANNHALTER, C., HINTERBERGER, W., HAAS, O., MAROSI, C., *et al.* (1991) Detection of engraftment and mixed chimerism following bone marrow transplantation using PCR amplification of a highly variable region- variable number of tandem repeats (VNTR) in the von Willebrand factor gene. *Ann Hematol* **63**: 227-228.

108. BADER, P., HOLLE, W., KLINGEBIEL, T., HANDGRETINGER, R., BENDA, N., *et al.* (1997) Mixed hematopoietic chimerism after allogeneic bone marrow transplantation: the impact of quantitative PCR analysis for prediction of relapse and graft rejection in children. *Bone Marrow Transplant* **19**: 697-702.

109. HIBI, S., TSUNAMOTO, K., TODO, S., SAWADA, T., UEDA, Y., *et al.* (1997) Chimerism analysis on mononuclear cells in the CSF after allogeneic bone marrow transplantation. *Bone Marrow Transplant* **20**: 503-506.

110. BADER, P., BECK, J., FREY, A., SCHLEGEL, P.G., HEBARTH, H., *et al.* (1998) Serial and quantitative analysis of mixed hematopoietic chimerism by PCR in patients with acute leukemias allows the prediction of relapse after allogeneic BMT. *Bone Marrow Transplant* **21**: 487-495.

111. ROSS, P., HALL, L. & HAFF, L.A. (2000) Quantitative approach to single-nucleotide polymorphism analysis using MALDI-TOF mass spectrometry. *Biotechniques* **29**: 620-626, 628-629.

112. LEUSHNER, J.R., WU, K., WANG, C., HOWSON, W. & METRAKOS, P. (1997) Temporal and quantitative analysis of chimerism in liver and kidney transplant patients: an Enhanced fluorescence detection system allows for more sensitive quantitative detection. *Mol Diagn* **2**: 187-195.

113. HAHN, A.B. & BALIGA, P. (2000) Rapid method for the analysis of peripheral chimerism in suspected graft-versus-host disease after liver transplantation. *Liver Transpl* **6**: 180-184.

114. KLEEBERGER, W., ROTHAMEL, T., GLOCKNER, S., LEHMANN, U. & KREIPE, H. (2000) Laser-assisted microdissection and short tandem repeat PCR for the investigation of graft chimerism after solid organ transplantation. *Pathobiology* **68**: 196-201.

115. GASSEL, H.J., OTTO, C., KLEIN, I., MEYER, D., TIMMERMANN, W., *et al.* (2000) Analysis of cellular events in hepatic allografts: donor progenitors induce intragraft chimerism. *Transpl Int* **13**: S465-470.

116. STARZL, T.E. (1993) Cell migration and chimerism - a unifying concept in transplantation - with particular reference to HLA matching and tolerance induction. *Transplant Proc* **25**: 8-12.

117. STARZL, T.E., DEMETRIS, A.J., MURASE, N., THOMSON, A.W., TRUCCO, M., *et al.* (1993) Donor cell chimerism permitted by immunosuppressive drugs: a new view of organ transplantation. *Immunol Today* **14**: 326-332.

118. STARZL, T.E., MURASE, N., DEMETRIS, A.J., ILDSTAD, S., RICORDI, C., *et al.* (1993) Allograft and xenograft acceptance under FK-506 and other immunosuppressant treatment. *Ann N Y Acad Sci* **685**: 46-51.

119. MAURY, S., JOUAULT, H., KUENTZ, M., VERNANT, J.P., TULLIEZ, M., *et al.* (2001) Chimerism analysis by lineage-specific fluorescent polymerase chain reaction in secondary graft failure after allogeneic stem cell transplantation. *Transplantation* **71**: 374-380.

120. ANDERSEN, C.B., LADEFOGED, S.D. & LARSEN, S. (1991) Cellular inflammatory infiltrates and renal cell turnover in kidney allografts: a study using *in situ* hybridization and combined *in situ* hybridization and immunohistochemistry with a Y-chromosome-specific DNA probe and monoclonal antibodies. *Apmis* **99**: 645-652.

121. ZHANG, J., TONG, K.L., LI, P.K., CHAN, A.Y., YEUNG, C.K., *et al.* (1999) Presence of donor- and recipient-derived DNA in cell-free urine samples of renal transplantation recipients: urinary DNA chimerism. *Clin Chem* **45**: 1741-1746.

122. SASANAKUL, W., HONGENG, S., CHUANSUMRIT, A., CHAIYARATANA, W., PAKAKASAMA, S., *et al.* (2001) The usefulness of X-linked polymorphic loci as gene markers to track X allele and chimerism in a post-allogeneic peripheral blood stem cell transplant patient with Wiskott-Aldrich syndrome. *J Med Assoc Thai* **84**: 379-384.

123. FEHSE, B., CHUKHLOVIN, A., KUHLCKE, K., MARINETZ, O., VORWIG, O., *et al.* (2001) Real-time quantitative Y-chromosome-specific PCR (QYCS-PCR) for monitoring hematopoietic chimerism after sex-mismatched allogeneic stem cell transplantation. *J Hematother Stem Cell Res* **10**: 419-425.

124. THIEDE, C., FLOREK, M., BORNHAUSER, M., RITTER, M., MOHR, B., *et al.* (1999) Rapid quantification of mixed chimerism using multiplex amplification of short tandem repeat markers and fluorescence detection. *Bone Marrow Transplant* **23**: 1055-1060.

125. JONE, C.M., AKEL, N. & KILLEEN, A.A. (2000) Evaluation of chimerism in DNA samples by PCR amplification of D1S80 with detection by capillary electrophoresis. *Mol Diagn* **5**: 101-105.

126. NOLLET, F., BILLIET, J., SELLESLAG, D. & CRIEL, A. (2001) Standardisation of multiplex fluorescent short tandem repeat analysis for chimerism testing. *Bone Marrow Transplant* **28**: 511-518.

127. KLEEBERGER, W., ROTHAMEL, T., GLOCKNER, S., FLEMMING, P., LEHMANN, U., *et al.* (2002) High frequency of epithelial chimerism in liver transplants demonstrated by microdissection and STR-analysis. *Hepatology* **35**: 110-116.

128. MUNIZ, E.S., PLASSA, F., AMSELEM, S., GOOSSENS, M. & VERNANT, J.P. (1994) Molecular analysis of polymorphic loci to study chimerism after allogeneic bone marrow transplantation. Heteroduplex analysis in denaturing gradient gel electrophoresis: a new approach to detecting residual host cells. *Transplantation* **57**: 451-456.

129. FORMANKOVA, R., HONZATKOVA, L., MORAVCOVA, J., SIEGLOVA, Z., DVORAKOVA, R., *et al.* (2000) Prediction and reversion of post-transplant relapse in patients with chronic myeloid leukemia using mixed chimerism and residual disease detection and adoptive immunotherapy. *Leuk Res* **24**: 339-347.

130. MILLSON, A.S., SPANGLER, F.L., WITTWER, C.T. & LYON, E. (2000) Comparison of automated short tandem repeat and manual variable number of tandem repeat analysis of chimerism in bone marrow transplant patients. *Diagn Mol Pathol* **9**: 91-97.

131. FINDLAY, I., RAY, P., QUIRKE, P., RUTHERFORD, A. & LILFORD, R. (1995) Allelic drop-out and preferential amplification in single cells and human blastomeres: implications for preimplantation diagnosis of sex and cystic fibrosis. *Hum Reprod* **10**: 1609-1618.

132. HESKETH, R. (1995) *The Oncogene Facts Book.*, Academic Press Limited, London.

133. YARDEN, Y. & SLIWKOWSKI, M.X. (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* **2**: 127-137.

134. MERIC, F., HUNG, M.C., HORTOBAGYI, G.N. & HUNT, K.K. (2002) HER2/neu in the management of invasive breast cancer. *J Am Coll Surg* **194**: 488-501.

135. BANKFALVI, A., SIMON, R., BRANDT, B., BURGER, H., VOLLMER, I., *et al.* (2000) Comparative methodological analysis of *erbB-2/HER-2* gene dosage, chromosomal copy number and protein overexpression in breast carcinoma tissues for diagnostic use. *Histopathology* **37**: 411-419.

136. SLAMON, D.J., CLARK, G.M., WONG, S.G., LEVIN, W.J., ULLRICH, A., *et al.* (1987) Human breast cancer: correlation of relapse and survival with amplification of the *HER-2/neu* oncogene. *Science* **235**: 177-182.

137. MUSS, H.B., THOR, A.D., BERRY, D.A., KUTE, T., T., L.E., *et al.* (1994) c-erbB-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. *N Engl J Med* **330**: 1260-1266.

138. BASELGA, J., NORTON, L., ALBANELL, J., KIM, Y.M. & MENDELSOHN, J. (1998) Recombinant humanized anti-HER2 antibody (Herceptin) enhances the antitumor activity of paclitaxel and doxorubicin against HER2/neu overexpressing human breast cancer xenografts. *Cancer Res* **58**: 2825-2831.

139. BASELGA, J., TRIPATHY, D., MENDELSOHN, J., BAUGHMAN, S., BENZ, C.C., *et al.* (1999) Phase II study of weekly intravenous trastuzumab (Herceptin) in patients with HER2/neu-overexpressing metastatic breast cancer. *Semin Oncol* **26**: 78-83.

140. ROCHE, P.C. & INGLE, J.N. (1999) Increased HER2 with U.S. Food and Drug Administration-approved antibody. *J Clin Oncol* **17**: 434-435.

141. PAULETTI, G., GODOLPHIN, W., PRESS, M.F. & SLAMON, D.J. (1996) Detection and quantification of *HER-2/neu* gene amplification in human breast cancer archival material using fluorescence *in situ* hybridization. *Oncogene* **13**: 63-72.

142. REVILLION, F., HORNEZ, L. & PEYRAT, J.P. (1997) Quantification of *c-erbB-2* gene expression in breast cancer by competitive RT-PCR. *Clin Chem* **43**: 2114-2120.

143. BIECHE, I., OLIVI, M., CHAMPEME, M.H., VIDAUD, D., LIDEREAU, R., *et al.* (1998) Novel approach to quantitative polymerase chain reaction using real-time detection: application to the detection of gene amplification in breast cancer. *Int J Cancer* **78**: 661-666.

144. BIECHE, I., ONODY, P., LAURENDEAU, I., OLIVI, M., VIDAUD, D., *et al.* (1999) Real-time reverse transcription-PCR assay for future management of ERBB2-based clinical applications. *Clin Chem* **45**: 1148-1156.

145. OKUYAMA, N., HATANO, Y., PARK, Y., SHIMATANI, S., SASAMOTO, S., *et al.* (1999) Quantification of *c-erbB-2* gene amplification in breast cancer tissue by competitive PCR. *Tumour Biol* **20**: 153-161.

146. LEHMANN, U., GLOCKNER, S., KLEEBERGER, W., VON WASIELEWSKI, H.F. & KREIPE, H. (2000) Detection of gene amplification in archival breast cancer specimens by laser-assisted microdissection and quantitative real-time polymerase chain reaction. *Am J Pathol* **156**: 1855-1864.

147. LYON, E., MILLSON, A., LOWERY, M.C., WOODS, R. & WITTWER, C.T. (2001) Quantification of *HER2/neu* gene amplification by competitive PCR using fluorescent melting curve analysis. *Clin Chem* **47**: 844-851.

148. SPECHT, K., RICHTER, T., MULLER, U., WALCH, A., WERNER, M., *et al.* (2001) Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol* **158**: 419-429.

149. ANDERS, F. (1967) Tumour formation in platyfish-swordtail hybrids as a problem of gene regulation. *Experientia* **23**: 10

150. KNUDSON JR., A.G. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA* **68**: 820-823.

151. PASSARGE, E. (1994) *Taschenatlas der Genetik.*, Georg Thieme Verlag, Stuttgart, New York.

152. CAVENEE, W.K., DRYJA, T.P., PHILLIPS, R.A., BENEDICT, W.F., GODBOUT, R., *et al.* (1983) Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* **305**: 779-784.

153. CAWKWELL, L., BELL, S.M., LEWIS, F.A., DIXON, M.F., TAYLOR, G.R., *et al.* (1993) Rapid detection of allel loss in colorectal tumours using microsatellites and fluorescent DNA technology. *Br J Cancer* **67**: 1262-1267.

154. LEVINE, A.J. & MOMAND, J. (1990) Tumor suppressor genes: the p53 and retinoblastoma sensitivity genes and gene products. *Biochim Biophys Acta* **1032**: 119-136.

155. HOLLSTEIN, M., SIDRANSKY, D., VOGELSTEIN, B. & HARRIS, C.C. (1991) p53 mutations in human cancer. *Science* **253**: 49-53.

156. HARRIS, C.C. & HOLLSTEIN, M. (1993) Clinical implications of the *p53* tumor-suppressor gene. *New Engl J Med* **392**: 1318-1327.

157. GREENBLATT, M.S., BENNET, W.P., HOLLSTEIN, M. & HARRIS, C.C. (1994) Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res* **54**: 4485-4478.

158. LANE, D.P. & CRAWFORD, L.V. (1979) T antigen is bound to a host protein in SV40-transformed cells. *Nature* **278**: 261-263.

159. LINZER, D.J., MALTZMAN, W. & LEVINE, A.J. (1979) The SV40 gene product is required for the production of a 54.000 MW cellular tumor antigen. *Virology* **98**: 308-318.

160. MCCORMIC, F. & HARLOW, E. (1980) Assoziation of a murine 53.000-dalton phosphoprotein with simian virus 40 large-T antigen in transformed cells. *J Virol* **34**: 213-224.

161. FINLAY, C.A., HINDS, P.W. & LEVINE, A.J. (1989) The p53 proto-oncogene can act as a suppressor of transformation. *Cell* **57**: 1083-1093.

162. LINVINGSTONE, L.R., WHITE, A., SPROUSE, J., LIVANOS, E., JACKS, T., *et al.* (1992) Altered cell cycle and gene amplification potential accompany loss of wild-type p53. *Cell* **70**: 923-935.

163. KASTAN, M.B., ZHAN, Q., EL-DEIRY, W.S., CARRIER, F., JACKS, T., *et al.* (1992) A mammalian cell cycle checkpoint utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell* **71**: 587-597.

164. KUERBITZ, S.J., PLUNKETT, B.S., WALSH, W.V. & KASTAN, M.B. (1992) Wild-type p53 is a cell cycle checkpoint determinant following irradiation. *Proc Natl Acad Sci USA* **89**: 7491-7495.

165. CAELLES, C., HELMBERG, A. & KARIN, M. (1994) p53-dependent apoptosis in the absence of transcriptional activation of p53-target genes. *Nature* **370**: 220-223.

166. YONISH-ROUACH, E., RESNITZKY, D., LOTEM, J., SACHS, L., KIMICHI, A., *et al.* (1991) Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6. *Nature* **352**: 345-347.

167. LIN, D., SHIELDS, M.T., ULLRICH, S.J., APPELLA, E. & MERCER, W.E. (1992) Growth arrest induced by wild-type p53 protein blocks cells prior or near the restriction point in late G1 phase. *Proc Natl Acad Sci USA* **89**: 9210-9214.

168. HARTWELL, L. (1992) Defects in a cell cycle checkpoint may be responsible for the genomic instability of cancer cells. *Cell* **71**: 543-546.

169. LANE, D.P. (1992) Worrying about p53. *Curr Biol* **2**: 581-583.

170. LEE, J.M. & BERNSTEIN, A. (1993) p53 mutations increase resistance to ionizing radiation. *Proc Natl Acad Sci USA* **90**: 5742-5746.

171. LOWE, S.W., RULEY, H.E., JACKS, T. & HOUSMAN, D.E. (1993) p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell* **74**: 957-967.

172. GRAEBER, T.G., OSMANIAN, C., JACKS, T., HOUSMAN, D.E., KOCH, C.J., *et al.* (1996) Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumors. *Nature* **379**: 88-91.

173. CALLAHAN, R. (1992) p53 mutations, another breast cancer prognostic factor. *J Natl Cancer Inst* **84**: 826-827.

174. HAMELIN, R., LAURENT-PUIG, P., OLSCHWANG, S., JEGO, N., ASSELAIN, B., *et al.* (1994) Association of p53 mutations with short survival in colorectal cancer. *Gastroenterol* **106**: 42-48.

175. DÖHNER, H., FISCHER, K., BENTZ, M., HANSEN, K., BENNER, A., *et al.* (1995) *p53* gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. *Blood* **85**: 1580-1589.

176. LEE, C.S., PIRDAS, A. & LEE, M.W. (1995) p53 in cutaneous melanoma: immunoreactivity and correlation with prognosis. *Australas J Dermatol* **36**: 192-195.

177. SERTH, J., KUCZYK, M.A., BOKEMEYER, C., HERVATIN, C., NAFE, R., *et al.* (1995) p53 immunohistochemistry as an independent prognostic factor for superficial transitional cell carcinoma of the bladder. *Br J Cancer* **71**: 201-205.

178. FINNEGAN, D.J. (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet* **5**: 103-107.

179. PATARCA, R. (1998) Chronic retroviruses and oncogenesis. *Crit Rev Oncogene* **9**: 151-198.

180. MCCLINTOCK, B. (1951) Chromosome organization and genic expression. *Cold Spring Harbor Symp Quant Biol* **16**,: 13-47.

181. COMFORT, N.C. (1995) Two genes, no enzyme: a second look at Barbara McClintock and the 1951 Cold Spring Harbour Symposium. *Genetics* **140**: 1161-1166.

182. FEDOROFF, N. & BOTSTEIN, D. (1992) Introduction. In: *The dynamic genome: Barbara McClintock´s ideas in the century of genetics* (FEDOROFF, N. & BOTSTEIN, D., eds).: 1-4. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

183. DEININGER, P.L., BATZER, M.A., HUTCHINSON, C.A. & EDGELL, M.H. (1992) Master genes in mammalian repetitive DNA amplification. *Trends Genet* **8**: 307-311.

184. CRAIG, N.L. (1997) Target site selection in transposition. *Ann Rev Biochem* **66**: 437 374

185. KRÖGER, M. & HOBOM, G. (1982) Structural analysis of insertion sequence IS5. *Nature* **297**: 159-162.

186. GALAS, D.J. & CHANDLER, M. (1989) Bacterial insertion sequences. In: *Mobile DNA* (BERG, D.E. & HOWE, M., eds).: 109-162. American Society of Microbiology, Washington DC.

187. BERG, D.E. (1989) Transposon Tn5. In: *Mobile DNA* (BERG, D.E. & HOWE, M., eds).: 185-210. American Society of Microbiology, Washington DC.

188. PLASTERK, R.H.A., IZSVÁK, Z. & IVICS, Z. (1999) Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet* **15**: 326-332.

189. RIO, D.C. (1991) Regulation of *Drosophila* P element transposition. *Trends Genet* **7**: 282-287.

190. ENGELS, W.R. (1997) Invasions of P elements. *Genetics* **145**: 11-15.

191. WEINER, A.M., DEININGER, P.L. & EFSTRATIADIS, A. (1982) Nonviral retroposons genes, pseudogenes and transposable elements generated by the reverse flow of genetic information. *Ann Rev Biochem* **55**: 631-661.

192. ONO, M., KAWAMAKI, M. & TAKEZAWA, T. (1987) A novel human nonviral retroposon derived from an endogenous retrovirus. *Nucleic Acids Res* **15**: 8725-8737.

193. KARYA, Y., KATO, K., HAYASHIZAKI, Y., HIMENO, S., TARUI, S., *et al.* (1987) Revision of consensus sequence of human Alu repeats. *Gene* **53**: 1-10.

194. MITSIALIS, S.A., MANLEY, J.L. & GUNTAKA, R.V. (1983) Localization of active promoters for eucaryotic RNA polymerase II in the long terminal repeat of avian sarcoma virus DNA. *Mol Cell Biol* **3**: 811-818.

195. FANNING, T.G. & SINGER, M.F. (1987) LINE-1: a mammalian transposable element. *Biochim Biophys Acta* **910**: 203-212.

196. BRATTHAUER, G.L. & FANNING, T.G. (1992) Active LINE-1 retrotransposons in human testicular cancer. *Oncogene* **7**: 507-510.

197. BRATTHAUER, G.L., CARDIFF, R.D. & FANNING, T.G. (1994) Expression of LINE-1 retrotransposons in human breast cancer. *Cancer* **173**: 2333-2336.

198. KATZ, R.A. & SKALKA, A.M. (1994) The retroviral enzymes. *Ann Rev Biochem* **63**: 133-173.

199. VARMUS, H.E. (1982) Form and function of retroviral provirus. *Science* **216**: 812-820.

200. LÖWER, R., LÖWER, J. & KURTH, R. (1996) The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc Natl Acad Sci USA* **93**: 5177-5184.

201. MICHIKAWA, Y., HOFHAUS, G., LERMAN, L.S. & ATTARDI, G. (1997) Comprehensive, rapid and sensitive detection of sequence variants of human mitochondrial tRNA genes. *Nucleic Acids Res* **25**: 2455-2463.

202. FRIEDHOFF, P. (1992) Quantitative PCR mit nachfolgender OLA/ELISA-Analytik. *Diplomarbeit*. Institut für Biochemie, Universität Hannover, Hannover.

203. MICHEL, J. (1999) Retrotransposons in der Antizipazion von Neoplasien: Untersuchungen am *Xiphophorus*-Modell der Melanombildung. *Inauguraldissertation*. Institut für Genetik, Justus-Liebig-Universität Gießen, Gießen.

204. FREIST, A. (2000) Entwicklungen von PCR-Methoden zur sensitiven Quantifizierung des mit der chronisch myeloischen Leukämie assoziierten BCR/ABL-Fusionstranskriptes. *Inauguraldissertation*. Institut für Biochemie, Justus-Liebig-Universität Gießen, Gießen.

205. HAHN, M. (1996) Entwicklung neuer PCR-Analyseverfahren für das humane Tumorsuppressorgen *p53*. *Inauguraldissertation*. Institut für Biochemie, Justus-Liebig-Universität Gießen, Gießen.

206. WITTWER, C.T., FILLMORE, G.C. & HILLIARD, D.R. (1989) Automated polymerase chain reaction in capillary tubes with hot air. *Nucleic Acids Res* **17**: 4353-4357.

207. AUSUBEL, F.M., BRENT, R., KINGSTON, R.E., MOORE, D.D., SEIDMAN, J.G., *et al.* (2001) *Current protocols in molecular biology.* John Wiley & Sons.

208. WITTWER, C.T., RIRIE, K.M., ANDREW, R.V., DAVID, D.A., GUNDRY, R.A., *et al.* (1997) The LightCycler(TM): A microvolume multisample fluorimeter with rapid temperature control. *Biotechniques* **22**: 176-181.

209. MANLY, B. (1997) *Randomization Bootstrap & Monte Carlo Methods in Biology*. Chapman & Hall, London.

210. KNUTH, D.E. (1981) Seminumerical Algorithms. In: *The Art of Computer Programming* Vol. **2**: 116ff. Edison-Wesley.

211. LONGLEY, M.J., BENNETT, S.E. & MOSBAUGH, D.W. (1990) Characterization of the 5' to 3' exonuclease associated with *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res* **18**: 7317-7322.

212. LYAMICHEV, V., BROW, M.A. & DAHLBERG, J.E. (1993) Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science* **260**: 778-783.

213. PECCOUD, J. & JACOB, C. (1996) Theoretical uncertainty of measurements using quantitative polymerase chain reaction. *Biophys J* **71**: 101-108.

214. CRAVEN, P. & WAHLBA, G. (1979) Smoothing noisy data with spline functions. *Numerische Mathematik* **31**: 377-403

215. DE BOOR, C. (2001) Calculation of the smoothing spline with weighted roughness measure. *Math Models Meth Appl Sci* **11**: 33-41

216. PINGOUD, A., URBANKE, C., HOGGET, J. & JELTSCH, A. (2002) *Biochemical Methods*. Wiley-VCH, Weinheim.

217. AL-ROBAIY, S., RUPF, S. & ESCHRICH, K. (2001) Rapid competitive PCR using melting curve analysis for DNA quantification. *Biotechniques* **31**: 1382-1386, 1388.

218. PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A. & VETTERLING, W.T. (1996) *Numerical recipes in Pascal*. Cambridge University Press, New York.

219. BERNARD, P.S., PRITHAM, G.H. & WITTWER, C.T. (1999) Color multiplexing hybridization probes using the apolipoprotein E locus as a model system for genotyping. *Anal Biochem* **273**: 221-228.

220. MARQUARDT, D.W. (1963) An algorithm for least-square estimation for non-linear parameters. *J Soc Ind Appl Math* **11**: 431-441.

221. GORDON, M. (1928) Pigment inheritance in the Mexican killifish: Interaction factors in *Platypoecilus maculatus*. *J Hered* **19**: 551-556.

222. KOSSWIG, C. (1928) Über Kreuzungen zwischen den Teleostiern *Xiphophorus helleri* und *Platypoecilus maculatus*. *Z. indukt. Abstammungs- und Vererbungslehre* **47**: 150-158.

223. HÄUSSLER, G. (1928) Über Melanombildungen bei *Xiphophorus helleri* und *Platypoecilus maculatus var. rubra*. *Klin. Wochenschrift* **7**: 1561-1562.

224. ANDERS, F., ANDERS, A. & VIELKIND, U. (1974) Regulation of tumor expression in the Gordon-Kosswig melanoma system, and the origin of malignancy. *XI. International Cancer Congress*, Florence: Panels 3, 305-306.

225. ANDERS, F., SCHARTL, M., BARNEKOW, A. & ANDERS, A. (1984) *Xiphophorus* as an *in vivo* model for studies on oncogenes. *Advances Cancer Res* **42**: 191-275.

226. ANDERS, A., ZECHEL, C., SCHLATTERER, B., GRÖGER, H., SCHMIDT, D., *et al.* (1991) Genetic and molecular approach for breeding and use of laboratory fish for the detection of agents with carcinogenic and/or promoting activity. *Bull Cancer* **78**: 415-433.

227. ROUSHDY, J., MICHEL, J., PETRY, H., ANDERS, A. & ANDERS, F. (1999) Paragenetic suppressors of suppressor genes - a new class of oncodeterminants. *J Cancer Res Clin Oncol* **125**: 123-133.

228. MITELMAN, F., JOHANSSON, B. & MERTENS, F. (2002) *Mitelman Database of Chromosome Aberrations in Cancer*. http://cgap.nci.nih.gov/Chromosomes/Mitelman.

229. ROETGER, A., BRANDT, B. & BARNEKOW, A. (1997) Competitive-differential polymerase chain reaction for gene dosage estimation of erbB-1 (*egfr*), *erbB-2*, and *erbB-3* oncogenes. *DNA Cell Biol* **16**: 443-448.

230. HYNES, N.E., GERBER, H.A., SAURER, S. & GRONER, B. (1989) Overexpression of the c-erbB-2 protein in human breast tumor cell lines. *J Cell Biochem* **39**: 167-173.

231. KRAUS, M.H., POPESCU, N.C., AMSBAUGH, S.C. & KING, C.R. (1987) Overexpression of the EGF receptor-related proto-oncogene *erbB-2* in human mammary cell lines by different molecular mechanisms. *EMBO J* **6**: 605-610.

232. FRYE, R.A., BENZ, C.C. & LIU, E. (1989) Detection of amplified oncogenes by differential polymerase chain reaction. *Oncogene* **4**: 1153-1157.

233. FEARON, E.R. & VOGELSTEIN, B. (1990) A genetic model for colorectal tumorigenesis. *Cell* **61**: 759-767.

234. NIGRO, J.M., BAKER, S.J., PREISINGER, A.C., JESSUP, J.M., HOSTETTER, R., *et al.* (1989) Mutations in the *p53* gene occur in diverse human tumour types. *Nature* **342**: 705-708.

235. FRIEND, S. (1994) p53: a glimpse at the puppet behind the shadow play. *Science* **265**: 334-335.

236. HAHN, M., MATZEN, S.E., SERTH, J. & PINGOUD, A. (1995) Semiautomated quantitative detection of loss of heterozygosity in the tumor suppressor gene *p53*. *BioTechniques* **18**: 1040-1047.

237. SHAPIRO, H.S. (1970) In: *CRC handbook of biochemistry. Selected data for molecular biology* (SOBER, H.A., ed): H-113. CRC Press, Cleveland, Ohio, USA.

238. COX, G.S., PHARES, C.K. & SCHMIDT, R.A. (1990) Molecular characterization of the *Spirometra mansonoides* genome: renaturation kinetics, methylation, and hybridization to human cDNA probes. *Biochem Biophys Acta* **1049**: 134-144.

239. SESHADRI, M. & RANJEKAR, P.K. (1980) Denaturation and reassoziation properties of the genome of *Phaseolus vulgaris*. *Hoppe-Seylers Z Physiol Chem* **361**: 1041-1048.

240. TIERSCH, T.R., CHANDLER, R.W., KALLMAN, K.D. & WACHTEL, S.S. (1989) Estimation of nuclear DNA content by flow cytometry in fishes of the genus *Xiphophorus*. *Comp Biochem Phys B* **94**: 465-468.

241. TIERSCH, T.R., CHANDLER, R.W., WACHTEL, S.S. & ELIAS, S. (1989) Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. *Cytometry* **10**: 706-710.

242. BINS, M. & TAKENS, F. (1985) A method to estimate the DNA content of whole nuclei from measurements made on thin tissue sections. *Cytometry* **6**: 234-237.

243. SUSNIK, B., POULIN, N., PHILLIPS, D., LERICHE, J. & PALCIC, B. (1995) Comparison of DNA measurement performed by flow and image cytometry of embedded breast tissue sections. *Anal Quant Cytol Histol* **17**: 163-171.

244. LEITCH, A.R., SCHWARZACHER, T., WANG, M.L., LEITCH, I.J., URLAN-MOMIRVIC, G., *et al.* (1992) Molecular cytogenetic analysis of repeated sequences in a long term wheed suspensions cultur. *Plant Cell Tissue Organ Culture* **33**: 287-296.

245. BINO, R.J., LANTER, S., VERHOEVEN, H.A. & KRAAK, H.L. (1993) Flow cytometic determination of nuclear replication stages in seed tissues. *Ann Bot* **72**: 181-187.

246. DOLEZEL, J., LUCRETTI, S. & SCHUBERT, I. (1994) Plant chromosome analysis and sorting by flow cytometry. *CRC Crit Rev Plant Sci* **13**: 275-309.

247. The genome sequence of *Sachharomyces cerevisiae* (1997) *Nature* **387**: 5-105.

248. BROWN, T.A. (1991) *Molecular Biology. LabFax.*, Bios Scientific Publishers, Oxfort, UK.

249. ALBERTS, B., BRAY, D., LEWIS, J., RAFF, M., ROBERTS, K., *et al.* (1994) *Molecular biology of the cell.*, Garland Publishing Group, New York, London.

250. KNIPPERS, R. (2001) *Molekulare Genetik.* Georg Thieme Verlag, Stuttgart.

251. SCHWAB, M. (1980) Sensitivität für Induktion von Neoplasien durch N-methyl-N-nitrosoharnstoff (MNH) als Folge der Kombination von Artgenomen bei *Xiphophorus*. *Habilitationsschrift*. Institut für Genetik, Justus-Liebig-Universität Gießen, Gießen.

252. LAY, M.J. & WITTWER, C.T. (1997) Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. *Clin Chem* **43**: 2262-2267.

253. RUIZ-PONTE, C., LOIDI, L., VEGA, A., CARRACEDO, A. & BARROS, F. (2000) Rapid real-time fluorescent PCR gene dosage test for the diagnosis of DNA duplications and deletions. *Clin Chem* **46**: 1574-1582.

254. GYLLENSTEN, U.B. & ERLICH, H.A. (1988) Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc Natl Acad Sci USA* **85**: 7652-7656.

255. PODDAR, S.K. (2000) Symmetric vs. asymmetric PCR and molecular beacon probe in the detection of a target gene of adenovirus. *Mol Cell Probes* **14**: 25-32.

256. SCHNELL, S. & MENDOZA, C. (1997) Theoretical description of the polymerase chain reaction. *J Theor Biol* **188**: 313-318.

257. GERMER, S., HOLLAND, M.J. & HIGUCHI, R. (2000) High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. *Genome Res* **10**: 258-266.

258. WU, D., UGOZZOLI, L., PAL, B.K. & WALLACE, R.B. (1989) Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. *Proc Natl Acad Sci USA* **86**: 2757-2760.

259. HUANG, M.M., ARNHEIM, N. & GOODMAN, M.F. (1992) Extension of base mispairs by *Taq* DNA polymerase: implications for single nucleotide discrimination in PCR. *Nucleic Acids Res* **20**: 4567-4573.

260. AYYADEVARA, S., THADEN, J.J. & SHMOOKLER REIS, R.J. (2000) Discrimination of primer 3'-nucleotide mismatch by *Taq* DNA polymerase during polymerase chain reaction. *Anal Biochem* **284**: 11-18.

261. MARCHETTI, G., PATRACCHINI, P., PAPACCHINI, M., FERRATI, M. & BERNARDI, F. (1993) A polymorphism in the 5' region of coagulation factor VII gene (*F7*) caused by an inserted decanucleotide. *Hum Genet* **90**: 575-576.

262. DAWSON, S.J., WIMAN, B. & HAMSTEN, A. (1993) The two allele sequences of a common polymorphism in the promotor of the plasminogen activator inhibitor-1 (*PAI-1*) gene respond differenfly to interleukin-l in HepG2 cells. *J Biol Chem* **268**: 10739-10745.

263. HUMPHRIES, S., TEMPLE, A., LANE, A., GREEN, F., COOPER, J., *et al.* (1996) Low plasma levels of factor VIIc and antigen are more strongly associated with the 10 base pair promoter (-323) insertion than the glutamine 353 variant. *Thromb Haemost* **75**: 567-572.

264. ERIKSSON, P., KALLIN, B., VAN'T HOFFI, F.M., BAVHENOLM, P. & HAMSTEN, A. (1995) Allele-specific increase in basal transcription of the plasminogen activator inhibitor-1 gene is associated with myocardial infarction. *Proc Natl Acad Sci USA* **92**: 1851-1855.

265. DI CASTELNUOVO, A., D'ORAZIO, A., AMORE, C., FALANGA, A., DONATI, M.B., *et al.* (2000) The decanukleotide insertion/deletion polymorphism in the promotor region of the coagulation factor VII gene and the risk of familial myocardial infarction. *Thromb Res* **98**: 9-17.

266. YE, S., GREEN, F.R., SCARABIN, P.Y., NICAUD, V., BARA, L., *et al.* (1995) The 4G/5G genetic polymorphism in the promoter of the plasminogen activator inhibitor-1 (*PAI-1*) gene is associated with differences in plasma PAI-1 activity but not with risk of myocardial infarction in the ECTIM study. *Thromb Haemost* **74**: 837-841.

267. SACCHI, E., TAGLIABUE, L., SCOGLIO, R., BARONCINI, C., COPPOLA, R., *et al.* (1996) Plasma factor VII levels are influenced by a polymorphism in the promotor region of the *FVII* gene. *Blood Coagul Fibrinolysis* **7**: 114-117.

268. GRENNET, H.E., KHAN, N., JIANG, W. & BOOYSE, F.M. (2000) Identification of the *Hin*dIII polymorphic site in the *PAI-1* gene: analysis of the *PAI Hin*dIII polymorphism by PCR. *Genet Test* **4**: 65-68.

269. HEZARD, N., CORNILLET, P., DROULLE, C., GILLOT, L., POTRON, G., *et al.* (1997) Factor V Leiden: detection in whole blood by ASA PCR using an additional mismatch in antepenultimate position. *Thromb Res* **88**: 59-66.

270. SASVARI-SZEKELY, M., GERSTER, A., RONAI, Z., STAUB, M. & GUTTMAN, A. (2000) Rapid genotyping of factor V Leiden mutation using single-tube bidirectional allele-specific amplification and automated ultrathin-layer agarose gel electrophoresis. *Electrophoresis* **21**: 816-821.

271. HAHN, M., WILHELM, J. & PINGOUD, A. (2001) Influence of fluorophor dye labels on the migration behaviour of polymerase chain reaction amplified short tandem repeats during denaturin capillary electrophoresis. *Electrophoresis* **22**: 2691-2700.

272. BERNARD, P.S. & WITTWER, C.T. (2002) Real-time PCR technology for cancer diagnostics. *Clin Chem* **48**: 1178-1185.

273. MACKAY, I.M., ARDEN, K.E. & NIETSCHE, A. (2002) Real-time PCR in virology. *Nucleic Acids Res* **30**: 1292-1305.

274. KANG, J.J., WATSON, R.M., FISHER, M.E., HIGUCHI, R., GELFAND, D.H., *et al.* (2000) Transcript quantitation in total yeast cellular RNA using kinetic PCR. *Nucleic Acids Res* **28**: e2.

275. BRODMANN, P.D., ILG, E.C., BERTHOUD, H. & HERRMANN, A. (2002) Real-time quantitative polymerase chain reaction methods for four genetically modified maize varieties and maize DNA content in food. *J AOAC Int* **85**: 646-653.

276. GIBSON, D.M. (1998) Rapid and definitive detection of Salmonella in foods by PCR. *Lett Appl Microbiol* **27**: 384-385.

277. HERNANDEZ, M., RIO, A., ESTEVE, T., PRAT, S. & PLA, M. (2001) A rapeseed-specific gene, acetyl-CoA carboxylase, can be used as a reference for qualitative and real-time quantitative PCR detection of transgenes from mixed food samples. *J Agric Food Chem* **49**: 3622-3627.

278. LOCATELLI, G., URSO, V. & MALNATI, M. (2000) Quantitative analysis of GMO food contaminations using real time PCR. *Ital J Biochem* **49**: 61-63.

279. NORTON, D.M. (2002) Polymerase chain reaction-based methods for detection of Listeria monocytogenes: toward real-time screening for food and environmental samples. *J AOAC Int* **85**: 505-515.

280. OVERBERGH, L., VALCKX, D., WAER, M. & MATHIEU, C. (1999) Quantification of murine cytokine mRNAs using real time quantitative reverse transcriptase PCR. *Cytokine* **11**: 305-312.

281. CHIANG, P.W., WEI, W.L., GIBSON, K., BODMER, R. & KURNIT, D.M. (1999) A fluorescent quantitative PCR approach to map gene deletions in the *Drosophila* genome. *Genetics* **153**: 1313-1316.

282. NIGRO, J.M., TAKAHASHI, M.A., GINZINGER, D.G., LAW, M., PASSE, S., *et al.* (2001) Detection of 1p and 19q loss in oligodendroglioma by quantitative microsatellite analysis, a real-time quantitative polymerase chain reaction assay. *Am J Pathol* **158**: 1253-1262.

283. MARCUCCI, G., LIVAK, K.J., BI, W., STROUT, M.P., BLOOMFIELD, C.D., *et al.* (1998) Detection of minimal residual disease in patients with AML1/ETO- associated acute myeloid leukemia using a novel quantitative reverse transcription polymerase chain reaction assay. *Leukemia* **12**: 1482-1489.

284. AMABILE, M., GIANNINI, B., TESTONI, N., MONTEFUSCO, V., ROSTI, G., *et al.* (2001) Real-time quantification of different types of bcr-abl transcript in chronic myeloid leukemia. *Haematologica* **86**: 252-259.

285. ELMAAGACLI, A.H., BEELEN, D.W., OPALKA, B., SEEBER, S. & SCHAEFER, U.W. (2000) The amount of BCR-ABL fusion transcripts detected by the real-time quantitative polymerase chain reaction method in patients with Philadelphia chromosome positive chronic myeloid leukemia correlates with the disease stage. *Ann Hematol* **79**: 424-431.

286. KRAUTER, J., HEIL, G. & GANSER, A. (2001) The AML1/MTG8 fusion transcript in t(8;21) positive AML and its implication for the detection of minimal residual disease; malignancy. *Hematol* **5**: 369-381

287. KRAUTER, J., HOELLGE, W., WATTJES, M.P., NAGEL, S., HEIDENREICH, O., *et al.* (2001) Detection and quantification of CBFB/MYH11 fusion transcripts in patients with inv(16)-positive acute myeloblastic leukemia by real-time RT-PCR. *Genes Chromosomes Cancer* **30**: 342-348.

288. MIYAZATO, A., UENO, S., OHMINE, K., UEDA, M., YOSHIDA, K., *et al.* (2001) Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hemato-poietic stem cell fraction. *Blood* **98**: 422-427.

289. RICKMAN, D.S., BOBEK, M.P., MISEK, D.E., KUICK, R., BLAIVAS, M., *et al.* (2001) Distinctive molecular profiles of high-grade and low-grade gliomas based on oligonucleotide microarray analysis. *Cancer Res* **61**: 6885-6891.

290. ROZZO, S.J., ALLARD, J.D., CHOUBEY, D., VYSE, T.J., IZUI, S., *et al.* (2001) Evidence for an interferon-inducible gene, Ifi202, in the susceptibility to systemic lupus. *Immunity* **15**: 435-443.

291. Frost & Sullivan Report B114 (2002).

292. GINZINGER, D.G. (2002) Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Haematol* **30**: 503-512.

293. MAKOTO, N., JANSSEN, B. & BARTRAM, C.R. (2001) Real-time detection of minimal residual desease by amplifiying immunoglobuline genes in acute lymphoblastic leukemia on the LightCycler. In: *Rapid cycle real-time PCR. Methods and applications.* (MEUER, S., WITTWER, C.T. & NAKAGAWARA, K., eds).: 197-206. Springer-Verlag, Heidelberg.

294. SCHMITT, Y. (2001) Performance characteristics of quantification assays for human immunodeficiency virus type 1 RNA. *J Clin Virol* **20**: 31-33.

295. TABERLET, P., GRIFFIN, S., GOOSSENS, B., QUESTIAU, S., MACANEAU, V., *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res* **24**: 3189-3194.

296. STENMAN, J. & ORPANA, A. (2001) Accuracy in amplification. *Nat Biotechnol* **19**: 1101-1102.

297. VON WURMB-SCHWARK, N., HIGUCHI, R., FENECH, A.P., ELFSTROEM, C., MEISSNER, C., *et al.* (2002) Quantification of human mitochondrial DNA in a real time PCR. *Forensic Sci Int* **126**: 34-39.

298. LEHMANN, U., LANGER, F., FEIST, H., GLOCKNER, S., HASEMEIER, B., *et al.* (2002) Quantitative assessment of promoter hypermethylation during breast cancer development. *Am J Pathol* **160**: 605-612.

299. GIULIETTI, A., OVERBERGH, L., VALCKX, D., DECALLONNE, B., BOUILLON, R., *et al.* (2001) An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. *Methods* **25**: 386-401.

300. TOYOOKA, K.O., TOYOOKA, S., MAITRA, A., FENG, Q., KIVIAT, N.C., *et al.* (2002) Establishment and validation of real-time polymerase chain reaction method for *CDH1* promoter methylation. *Am J Pathol* **161**: 629-634.

301. JANSSEN, B.L., VAN DE LOCHT, L.T., FOURKOUR, A., DE SMET, C., MENSINK, E.J., *et al.* (1999) Transcription of the MAGE-1 gene and the methylation status of its Ets binding promoter elements: a quantitative analysis in melanoma cell lines using a real-time polymerase chain reaction technique. *Melanoma Res* **9**: 213-222.

302. LO, Y.M., WONG, I.H., ZHANG, J., TEIN, M.S., NG, M.H., *et al.* (1999) Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction. *Cancer Res* **59**: 3899-3903.

303. MULLER-TIDOW, C., BORNEMANN, C., DIEDERICHS, S., WESTERMANN, A., KLUMPEN, S., *et al.* (2001) Analyses of the genomic methylation status of the human cyclin A1 promoter by a novel real-time PCR-based methodology. *FEBS Lett* **490**: 75-78.

304. GINZINGER, D.G., GODFREY, T.E., NIGRO, J., MOORE, D.H., 2ND, SUZUKI, S., *et al.* (2000) Measurement of DNA copy number at microsatellite loci using quantitative PCR analysis. *Cancer Res* **60**: 5405-5409.

305. MORGAN, K.T., NI, H., BROWN, H.R., YOON, L., QUALLS, C.W., JR., *et al.* (2002) Application of cDNA microarray technology to in vitro toxicology and the selection of genes for a real-time RT-PCR-based screen for oxidative stress in Hep-G2 cells. *Toxicol Pathol* **30**: 435-451.

306. FINK, L., SEEGER, W., ERMERT, L., HANZE, J., STAHL, U., *et al.* (1998) Real-time quantitative RT-PCR after laser-assisted cell picking. *Nat Med* **4**: 1329-1333.

307. GLOCKNER, S., LEHMANN, U., WILKE, N., KLEEBERGER, W., LANGER, F., *et al.* (2000) Detection of gene amplification in intraductal and infiltrating breast cancer by laser-assisted microdissection and quantitative real-time PCR. *Pathobiology* **68**: 173-179

308. EHRIG, T., ABDULKADIR, S.A., DINTZIS, S.M., MILBRANDT, J. & WATSON, M.A. (2001) Quantitative amplification of genomic DNA from histological tissue sections after staining with nuclear dyes and laser capture microdissection. *J Mol Diagn* **3**: 22-25.

309. GODFREY, T.E., KIM, S.H., CHAVIRA, M., RUFF, D.W., WARREN, R.S., *et al.* (2000) Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. *J Mol Diagn* **2**: 84-91.

310. LEHMANN, U., BOCK, O., GLOCKNER, S. & KREIPE, H. (2000) Quantitative molecular analysis of laser-microdissected paraffin- embedded human tissues. *Pathobiology* **68**: 202-208.

311. RICE, J. & ROSENBLATT, M. (1983) Smoothing Splines: regression, derivatives and deconvolution. *Annals of Statistics* **11**: 141-156.

7 ANHANG

7.1 Messung, Rauschen und Validierung

7.1.1 Messung

Ein Messvorgang liefert als Ergebnis den Wert einer Messgröße mit einer Einheit, die sich aus dem internationalen System der Einheiten (SI-System) ableiten lässt. Die Messgröße ist eine Eigenschaft des Messobjekts, welche nach einem definierten Verfahren in reproduzierbarer Weise von einem Messinstrument quantifiziert wird.

Prinzipiell beeinflusst jeder Messvorgang das Messobjekt und damit das Ergebnis der Messung. Die Beeinflussung ist aber meist vernachlässigbar klein. Nicht vernachlässigbare Beeinflussungen des Messobjekts durch den Messvorgang können zu systematischen Fehlern führen. Ist die Art und Stärke der Beeinflussung bekannt, kann der systematische Fehler korrigiert werden.

Jeder Messvorgang wird weiterhin beeinflusst durch unkontrollierbare Störungen auf das Messobjekt, den Messvorgang und das Messinstrument. Sie führen zu einer Streuung der Messwerte. Zur Ermittlung eines genauen Ergebnisses wird daher die Messung mehrfach durchgeführt und aus den Werten dieser Stichprobe mit statistischen Verfahren der Wertebereich bestimmt, der mit einer vorgegebenen Wahrscheinlichkeit den wahren Wert der Messgröße beinhaltet.

7.1.2 Rauschen

In der Mess- und Signaltechnik bezeichnet man die durch unkontrollierbare Einflüsse verursachten Streuungen der Messwerte als Rauschen. Rauschen beinhaltet alle jene Anteile am Signal, die keine brauchbare Information enthalten. Es sind daher durchweg unerwünschte Anteile. Andererseits lassen sich aus der Amplitudenverteilung, dem Frequenzspektrum, dem zeitlichen Verlauf und der Abhängigkeit äußerer Einflüssen auf die Rauschquelle neue Informationen gewinnen. Rauschen kann viele Ursachen haben. Einige wichtige sind hier mit ihren Eigenschaften aufgeführt.

1) **Thermisches Rauschen** (*Johnson Noise*). Durch die zufällige Bewegung von Ladungsträgern (Elektronen, Atome, Moleküle) in einem Leiter entstehen zufällige Spannungsschwankungen unterschiedlichster Frequenzen. Ein Widerstand, wichtiges Bauteil vieler elektronischer Messinstrumente, ist also eine Rauschspannungsquelle. Die Höhe der maximalen Rauschleistung, die diese Quelle abgeben kann, ist proportional zur absoluten Temperatur und in jedem Frequenzintervall fester Breite gleich groß (weißes Rauschen): $\Delta P = \Delta f k T$ (P: Rauschleistung, f: Frequenz in Hertz, k: Boltzmann-Konstante

$k = 1.38 \cdot 10^{-23}$ Ws/K, T: absolute Temperatur in Kelvin). Da die Gesamtleistung ($\Delta f = \infty$) nicht unendlich groß werden kann, ist dies nur eine Näherung. Die spektrale Leistungsdichte muss bei großen Frequenzen kleiner werden. Sie ist aber konstant in dem Frequenzbereich, in dem üblicherweise messtechnische Signale vorkommen. Bei Raumtemperatur nimmt sie nach dem Planckschen Strahlungsgesetz erst bei etwa 1GHz deutlich ab.

2) **Quantisierungsrauschen**. Bei der Digitalisierung eines analogen Signals können die Abweichungen zwischen dem Wert des analogen und des digitalen Signals bis zur Hälfte einer Stufenhöhe betragen. Das Quantisierungsrauschen enthält im Wesentlichen die Abtastfrequenz und deren Oberschwingungen.

3) **Netzbrummen**. Die wird bewirkt durch eine induktive oder kapazitive Ankopplung des elektrischen Versorgungsnetzes (50 Hz). Wenn hinreichend lange Zeitreihen von Einzelmessungen vorliegen, lässt sich der Einfluss des Netzbrummens durch eine Fourier-Analyse bestimmen und eliminieren.

7.1.3 Validierung

Wie kann man nun sicher sein, dass bei einer Messung wirklich die gesuchte Größe gemessen wurde? Wie nahe kommt der Messwert dem tatsächlichen Wert der Größe? Maßnahmen zur Lösung dieser Probleme nennt man Validierung. Dabei ermittelte Größen sind Validierungs-Parameter (*validation parameters*). Zur besseren Übersicht lassen sie sich einteilen in:

1) Parameter bezüglich eines Wertes

Diese Größen beschreiben die Beziehung eines oder mehrerer Messwerte zum zugehörigen, tatsächlichen (aber meist unbekannten) Wert.

Richtigkeit (*accuracy*). Mit dem Begriff „Richtigkeit" wird beschrieben, ob die Messwerte um den tatsächlichen, den richtigen bzw. wahren Wert streuen. Sind nur zufällige Schwankungen vorhanden, wird man erwarten, dass mit zunehmender Anzahl an Messungen der Mittelwert sich immer mehr dem richtigen Wert nähert. Der Abstand dieser beiden Werte (absolut oder relativ) ist die Richtigkeit. Eine hohe Richtigkeit kann mit großer oder kleiner Streuung (Präzision) erreicht werden.

Präzision (*precision*). Die Präzision ist ein Maß für die Streuung der einzelnen Messwerte um ihren Mittelwert. Zur Angabe wird häufig die Standardabweichung verwendet oder es wird ein Intervall angegeben, in dem der nächste Messwert mit einer bestimmten Wahrscheinlichkeit anzutreffen sein wird (z. B. 95%). Größere Wahrscheinlichkeiten führen dabei auch zu größeren Intervallen. In der Präzision steckt keine Aussage über die Richtigkeit. Die Präzision lässt sich bezüglich der folgenden Wiederholungsbedingungen unterscheiden:

i) Wiederholbarkeit (*repeatability*): ein Labor und ein Experimentator, kurzer Zeitabstand

ii) Reproduzierbarkeit (*reproducibility*): anderes Labor, Gerät oder Experimentator

Selektivität (*specificity*, *selectivity*). Manche Messergebnisse sind davon abhängig, ob z.B. andere Substanzen oder andere Verfahrensverläufe vorhanden sind oder sich die Werte anderer Größen geändert haben. Die Selektivität gibt an, wie stark das Ergebnis von anderen Einflüssen abhängig ist, ob also auch nur die zu messende Größe gemessen wird. Manchmal wird auch der Begriff "Querempfindlichkeit" verwendet.

2) *Parameter bezüglich eines Verfahrens*

Diese Größen beschreiben Möglichkeiten und Grenzen von Messungen nach einem bestimmten Verfahren.

Empfindlichkeit (*sensitivity*). Die Empfindlichkeit einer Messung ist durch die Nachweisgrenze sowie die Erfassungsgrenze gegeben (s.u.).

Auflösung (*resolution*). Ein Maß für die Auflösung einer Messung ist durch die Bestimmungsgrenze gegeben (s.u.).

Linearität (*linearity*). Die Linearität gibt an, wie sich die Genauigkeit und Präzision von Messungen in Abhängigkeit von dem Wert der Messgröße verhalten.

Hysterese (*hysteresis*). Von einer Hysterese spricht man, wenn der Endwert einer sich einstellenden Größe von der Richtung abhängt, aus der die Einstellung erfolgt. Dann ist die Zuordnung von Eingangs- und Ausgangsgröße eines Systems nicht mehr eindeutig. Hysterese tritt z. B. bei mechanischen Vorgängen mit Spiel (Getriebe, Zeiger-Anzeige), bei magnetischer Ummagnetisierung oder bei mechanischer Belastung im nichtelastischen Bereich auf.

Messbereich (*range*). Der Messbereich ist das Intervall für die Eingangsgröße, innerhalb derer bestimmte Eigenschaften (Fehler, Stabilität) garantiert werden. Der Anzeigebereich kann größer als der Messbereich sein. Mit I.D.L. (*Instrumental Device Limit*) wird die Grenze bezeichnet, die sich allein aus den Grenzen des Meßsystems ohne Berücksichtigung von Proben- oder Verfahrensunsicherheiten ergibt.

Nachweisgrenze (*limit of detection*). An der Nachweisgrenze (ein Wert der Merkmalsgröße) erhält man einen Wert der Signalgröße, der mit 50%iger Wahrscheinlichkeit von der zu messenden Größe stammt, genauso gut aber durch zufällige Schwankungen verursacht sein kann. Unterhalb der Nachweisgrenze sind weder qualitative noch quantitative Aussagen möglich.

Erfassungsgrenze (*limit of quantification*). Oberhalb der Erfassungsgrenze ist die Wahrscheinlichkeit für ein zufälliges Signal dieser Größe gering (z.B. kleiner als 1 %). Zwischen Nachweis- und Erfassungsgrenze kann qualitativ, oberhalb der Erfassungsgrenze quantitativ gemessen werden.

Bestimmungsgrenze. Quantitative Aussagen können auch oberhalb der Erfassungsgrenze nur mit einer gewissen Unsicherheit (Fehler) gemacht werden. Die Bestimmungsgrenze legt einen Merkmalswert fest, oberhalb dessen der Fehler kleiner als eine festgelegte Schranke (z.B. 5 %) ist.

Robustheit (*robustness, ruggedness*). Die Robustheit beschreibt bei Messprozessen die Unempfindlichkeit gegenüber Störeinflüssen, welche genau beschreibbar sein müssen. Bei Produktionsprozessen wird eine stabile Ausgangsgröße gefordert. Dort wird mit Robustheit auch die Unempfindlichkeit gegenüber Schwankungen der Eingangsgröße beschrieben.

7.2 Nichtparametrische *Splines*

Als *Splines* wurden die ursprünglich im Schiffsbau verwendeten flexiblen Holzplanken bezeichnet, die an bestimmten Punkten fixiert wurden, während sie sich dazwischen frei biegen durften. Im mathematischen Sinne handelt es sich bei *Splines* um eine Funktion $s(x)$, die in einem Intervall $(a;b)$ stückweise aus einfachen Polynomen zusammengesetzt ist, wobei eine vorgegebene Anzahl ihrer Ableitungen aller Punkte im Intervall $(a;b)$ kontinuierlich sind. Der Vergleich zu den Schiffsplanken ergibt sich insbesondere dadurch, dass die Polynome durch die gegebenen Stützpunkte oder Knoten (die zu interpolierenden Datenpunkte x_i/y_i) festgelegt sind und ihr Verlauf so angepasst wird, dass die Übergänge der Steigungen (s') und Krümmungen (s'') in den Stützpunkten wie bei einer gebogenen Holzplanke kontinuierlich sind und die Gesamtkrümmung minimal ist. Um eine Kontinuität der Krümmungen im *Spline* zu erreichen, müssen die zusammengesetzten Polynome mindestens dreimal differenzierbar, also dritten Grades, sein. Aus diesem Grund werden zumeist kubische *Splines* verwendet, die wie folgt definiert sind:

$$s(x) = a_i + b_i \cdot x + c_i \cdot x^2 + d_i \cdot x^3 \qquad \textbf{Gl. 7.1}$$

wobei x_i mit $i = 1 \ldots n$ die Lage der Knoten beschreibt und $x_i \leq x < x_{i+1}$ gilt.

Aus den Kontinuitätsbedingungen der Funktionswerte sowie der ersten beiden Ableitungen ergeben sich folgende Einschränkungen:

$$a_{i-1} + b_{i-1}x_i + c_{i-1}x_i^2 + d_{i-1}x_i^3 = a_i + b_i x_i + c_i x_i^2 + d_i x_i^3 \quad \text{(Funktionswerte)}$$

$$b_{i-1} + 2c_{i-1}x_i + 3d_{i-1}x_i^2 = b_i + 2c_i x_i + 3d_i x_i^2 \quad \text{(Steigungen bzw. 1.Ableitungen)}$$

$$2c_{i-1} + 6d_{i-1}x_i = 2c_i + 6d_i x_i \quad \text{(Krümmungen bzw. 2.Ablleitungen)}$$

Eine weitere Bedingung ist dadurch gegeben, dass das *Spline* eine definierte Gesamtkrümmung aufweisen soll. Dazu wird das Integral der quadrierten zweiten Ableitung vorgegeben:

$$\int (s'')^2 = w \qquad \textbf{Gl. 7.2}$$

Für den Grenzfall $w = w_{\max}$ geht das *Spline* zwingend durch die angegebenen Knoten, während sich für $w = 0$ die Regressionsgerade durch die Knotenpunkte ergibt. In diesem Fall wird die Summe der Abweichungsquadrate $\sum\left(y_i - s(x_i)\right)^2$ maximal. Ein *Spline*, welches die Summe der Abweichungsquadrate minimiert, dabei aber noch eine hinreichende Gesamtkrümmung aufweist, um einem nichtlinearen Datenverlauf lokal folgen zu können, lässt sich unter Einführung eines Glättungsparameters p durch folgende Bedingung ermitteln:

$$R(p) = p \cdot \sum\left(y_i - s(x_i)\right)^2 + (1-p) \cdot \int s''(x)^2 \qquad \textbf{Gl. 7.3}$$

Über den Parameter p wird so eine Verbindung hergestellt zwischen der Summe der Fehlerquadrate und der lokalen Krümmung. Diejenige *Spline*-Funktion $s(x)$, welche $R(p)$ für einen gegebenen Wert von p minimiert, wird als glättender kubischer *Spline* bezeichnet.

Eine genaue Herleitung der *Spline*-Funktion, das Lösungsverfahren von **Gl. 7.3** sowie die programmtechnischen Umsetzungen sind in verschiedenen Artikeln über Numerische Mathematik [214,215] und Statistik [311] ausführlich beschrieben und diskutiert.

7.3 Begriffsdefinitionen

Tab. 7-1 Begriffserklärungen zu Programmierung und RT-PCR-Experimenten.

Begriff	Erklärung
Datentyp	Der Datentyp bestimmt die Art und Weise, wie Daten im Speicher des Rechners abgelegt werden und vom Prozessor interpretiert werden. Wichtige fundamentale Datentypen sind u. a. Wahrheitswerte (Boolean), Ganzzahlen (Integer), Fließkommawerte (z. B. Single) und Zeichenketten (String). Pascal erlaubt die Definition eigener Datentypen. Konventionsgemäß beginnen die Typenbezeichner mit einem `T`, z. B. `TMeinTyp`.
Eigenschaft	Eigenschaften (*properties*) stellen Schnittstellen zu den Daten, z. B. den Feldern einer Klasse, dar. Die Werte von Eigenschaften können von der Klasse indirekt über den Aufruf entsprechender Methoden geändert oder bestimmt werden. Es können auch nur-lesen-Eigenschaften definiert werden, deren Werte von außerhalb des Objekts nicht geändert werden können.
Ereignis	Ein Ereignis (*event*) kann von einem Objekt ausgelöst werden, nachdem das Objekt eine Aktion durchgeführt hat. Dazu wird die zugewiesene Ereignisbehandlungsmethode aufgerufen. Andere Teile des Programms können auf das Ereignis reagieren, indem sie der Ereignis-Eigenschaft die Methode zuweisen, in welcher der dann auszuführende Code steht.
Experiment	Die Durchführung einer PCR mit dem LightCycler. Das Ergebnis eines Experiments sind die vom LightCycler erstellten Dateien.

Begriff	Erklärung
Feld	Felder sind die Datenelemente (Variablen) einer Klasse. Die Bezeichner von Feldvariablen beginnen konventionell mit einem `F`. Felder sollten außerhalb des betreffenden Objekts nur über Eigenschaften bzw. Methoden zugänglich sein, damit das Objekt selbst kontrollieren kann, ob die Änderungen sinnvoll sind.
Klasse	Eine Klasse (*class*) ist eine abstrakte Programmstruktur, welche den Inhalt und das Verhalten von Objekten festlegt. Die Klassen werden als Datentypen deklariert, die Bezeichner von Klassen beginnen also ebenfalls mit einem `T`. Neben den Deklarationen von Datenfeldern beinhalten Klassen auch Programmcode in Form sogenannter Methoden, die ihr Verhalten steuern. Der Zugriff auf die Inhalte der Datenfelder erfolgt über sogenannte Eigenschaften. Neue Klassen können von bereits programmierten Klassen abgeleitet werden. Die neue Klasse erbt dann alle Datenfelder, Methoden und Eigenschaften der Parentalklasse, die in der Neudeklaration verändert, erweitert und ergänzt werden können.
Komponente	Eine Komponente ist eine Klasse, die sich während der Programmentwicklung visuell in das Programm einbinden lässt. Sie leitet sich ab von der in Delphi vordefinierten Klasse `TComponent`. Die Eigenschaften und Ereignisse von Komponenten lassen sich während der Programmentwicklung im Objektinspektor von Delphi ändern und werden gemeinsam mit dem Projekt abgespeichert. Die Schaltflächen, Textanzeigen, und Menüs sind Beispiele für Komponenten.
Methode	Codeabschnitt (Prozedur bzw. Funktion) einer Klasse.
Objekt	Ein Objekt (*object*) ist eine Instanz bzw. eine Variable einer Klasse. Im Gegensatz zu Variablen anderer Datentypen (die keinen Programmcode enthalten), müssen Objekt-Variablen durch einen Aufruf des Klassenkonstruktors erzeugt werden. Der Konstruktor reserviert den benötigten Speicherplatz und erstellt eine Zugriffstabelle auf die Methoden.
Programm	Definition der Temperaturprofile und Messpunkte in einem LightCycler-Experiment. Ein Programm enthält Angaben über den Analysemodus der Messdaten (Quantifizierung, Schmelzkurve oder keine Analyse), die Zyklenzahl und die Temperatur- und Messpunkteinstellungen in Form von Segmenten. Der LightCycler erlaubt die Benennung von Programmen (z. B. „Denat", „Ampli", „Melt", usw.).
Protokoll	Gerätespezifische Angaben zum Experiment wie die Einstellungen der LED, Verstärkungsfaktoren der Photohybriden und die Programme zur Temperatur- und Messsteuerung.
Segment	Teil eines Programms mit Angaben über die Zieltemperatur, Heiz- bzw. Kühlraten, Dauer und Lage der Messpunkte. Es kann eine zweite Zieltemperatur angegeben werden, welche bei den zyklischen Wiederholungen durch definierte Temperaturinkremente stufenweise eingestellt wird. Die Art des Messpunkts wird als Acquisitionsmodus bezeichnet. Der LightCycler erlaubt die einmalige Aufzeichnung eines Messwerts am Ende des Segments (*single*), die kontinuierliche Aufzeichnung während des gesamten Segments (*continous*) sowie die schrittweise Aufzeichnung während der Temperaturänderung (*stepwise*). Nähere Informationen hierzu stehen im LightCycler-Benutzerhandbuch, Version 3.5, Seite 80.
Variable	Die Variable bezeichnet eine Adresse im Speicher des Rechners. Jede Variable hat einen definierten Datentyp, der bestimmt, wie die an dieser Adresse abgelegten

Begriff	Erklärung
	Daten vom Prozessor zu interpretieren sind.
virtuelle Methode	Überschreibbarer Codeabschnitt einer polymorphen Klasse. Definiert man z. B. eine Klasse `TShape` zur grafischen Darstellung einer geometrischen Figur mit der Methode Draw, welche die Figur auf einer Zeichenfläche ausgibt, kann in `TShape.Draw` noch nicht definiert sein und wird virtuell deklariert. Bei abgeleiteten Objekte wie `TCircle` oder `TSquare` kann nun die Methode `Draw` überschrieben werden, in diesem Beispiel mit Code zum Zeichnen eines Kreises oder eines Quadrats. Eine andere Prozedur zum Zeichnen komplexer Bilder könnte nun z. B. Objekte vom Typ `TShape` verwalten. Zum Zeichnen dieser Objekte ruft diese Prozedur die Methode Draw dieser Objekte auf als `TShape.Draw` ohne nähere Informationen darüber zu benötigen, ob es sich dabei um `TCircle` oder `TSquare` oder ein anderes, von `TShape` abgeleitetes Objekt handelt (was vielleicht zum Zeitpunkt der Programmierung noch gar nicht existiert). Da diese Methode virtuell ist, können alle von `TShape` abgeleiteten Objekte durch einen Aufruf von `TShape. Draw` gezeichnet werden. Wenn das `TShape`-Objekt, dessen Methode Draw aufgerufen wird, mit dem Konstruktor von `TCircle` erstellt wurde, wird in diesem Fall die Methode `TCircle.Draw` aufgerufen und ein Kreis gezeichnet.

7.4 Klassenbeschreibungen

Tab. 7-2 Kurzbeschreibung der Klassen zur Verwaltung und Auswertung von RT-PCR-Daten.

Klassenbezeichner	Beschreibung
`TAnalysis`	Abstrakte Klasse zur Durchführung von Analysen. Verwaltet Listen analysespezifischer Daten zu Proben und Probengruppen. Die Eigenschaften Segment und `ChannelIndex` bestimmen das Segment und den Kanal bzw. das Kanalverhältnis, dessen Messdaten analysiert werden sollen. Neuberechnungen über die Methode Update lösen das Ereignis `OnResultsChange` aus, auf welches das Hauptprogramm reagieren kann.
`TAnalysisData`	Verwaltet die zur Analyse benötigten x- und y-Daten einer Probe. Die Methode `CalcData` berechnet in Abhängigkeit der Werte von Segment und `ChannelIndex` der übergeordneten Analyse die x- und y-Daten.
`TAnalysisGroup`	Verwaltet die probenbezogenen Analysedaten von Proben, die zu einer gemeinsamen Gruppe gehören.
`TChnData`	Speichert eine Liste von Objekten des Typs `TFltList`.
`TDataList`	Speichert eine Liste von Objekten des Typs `TSegData`.
`TDataList`	Die von `TDataObjekt` abgeleitete Klasse ermöglicht die Definition der

Klassenbezeichner	Beschreibung
	Listengröße über die Eigenschaft `Count` und erlaubt mit den Methoden `Insert` und `Remove` das Einfügen und Entfernen einzelner Elemente aus der Datenliste.
`TDataObject`	Abstrakte Klasse zur Verwaltung von Datenlisten mit Elementen gleicher Größe. Die Daten von Instanzen dieser Klasse lassen sich gegenseitig zuweisen, hinzufügen, laden und speichern. Die Klasse erlaubt das Löschen der Datenwerte oder der Daten. Die Zuweisung des Speicherplatzes geschieht dynamisch. Der Zugriff auf die Daten erfolgt über die Eigenschaft `Data`, die einen Zeiger auf den Speicherbereich der Daten zurückgibt.
`TDataSegments`	Speichert eine Liste von Objekten des Typs `TSegList`, sortiert nach `AnalysisMode` der übergeordneten Programme. Wird von `TRun` verwaltet und erlaubt den direkten Zugriff auf diejenigen Segmente, die zu analysierende Messdaten enthalten.
`TFltList`	Speichert eine Liste von Fließkommazahlen.
`TGroup`	Identisch mit `TSmpList`.
`TGroups`	Speichert eine Liste von Objekten des Typs `TGroup`.
`TLCTreeview`	Verwaltet eine Liste von Objekten des Typs `TRun` sowie eine Liste mit Anzeigeparametern, zeigt die enthaltenen Proben in einer Baumstruktur an und erlaubt die Selektion von Proben. Die gleichzeitig verwalteten `TRun`-Objekte müssen das gleiche Protokoll aufweisen. Jede Änderungen innerhalb des Objekts löst ein entsprechendes Ereignis aus, worauf im Hauptprogramm reagiert werden kann.
`TListObj`	Die von dieser Klasse abgeleiteten Klassen können von `TObjList`-Nachfolgern verwaltet werden. Über die Eigenschaft `ObjList` ist ein Zugriff auf die Liste möglich, die Funktion `Index` gibt die Position dieses Elements in der übergeordneten Liste zurück. Mit den Funktionen `Next` und `Prev` lässt sich das jeweils nächste bzw. vorige Objekt in der Liste zurückgeben. Über die Eigenschaft `Text` lässt sich dem Objekt eine Zeichenkette zur Identifizierung zuweisen.
`TMeltAnalysis`	Komponentenklasse zur Durchführung von Schmelzkurvenanalysen. Über die Eigenschaft `QuenchCorrection` wird bestimmt, ob eine Korrektur des temperaturbedingten Quenchs durchgeführt wird. Mit der Eigenschaft `Threshold` lässt sich die Lage des Schwellenwerts zur Peakdetektion definieren. Die Eigenschaft `AutoAdjust` legt fest, ob der Schwellenwert automatisch anhand des mittleren Signalrauschens bestimmt wird.
`TMeltData`	Verwaltet die zur Schmelzkurvenanalyse benötigten Daten einer Probe. Führt die Korrektur des temperaturbedingten Quenchs durch und berechnet die Peakdaten. Die Eigenschaft `PeakData` gibt die negative erste Ableitung der Schmelzkurvendaten zurück, über die Eigenschaft Peaks ist der Zugriff auf die Parameter der bestimmten Schmelzpeaks möglich.

Klassenbezeichner	Beschreibung
TMeltGroup	Identisch mit TAnalysisGroup.
TObjList	Verwaltet eine Liste von Objekten des Typs TListObj. Die Liste kann die enthaltenen Objekte automatisch löschen, wenn die Liste gelöscht wird. Abgeleitete Klassen verwalten entsprechend Objekte, die von TListObj abgeleitet sind.
TPrgList	Speichert eine Liste von Objekten des Typs TProgram.
TProgram	Speichert die Daten eines LightCycler-Programms.
TQuantAnalysis	Komponentenklasse zur Durchführung von Quantifizierungen. Berechnet u. a. eine Kalibriergerade anhand der C_T-Werte selektierter Standards und ermittelt den Status der Quantifizierung. Die Eigenschaften BgrMode, AmpFunc und CTMode bestimmen die Art der Hintergrundkorrektur, die verwendete Amplifikationsfunktion zur Bestimmung der C_T-Werte und das Verfahren zur Berechnung der C_T-Werte. Über die Eigenschaften *Noiseband* und NBPoints können die Lage des Schwellenwerts und die Anzahl verwendeter Punkte für die C_T-Wert-Bestimmung nach der *Fit-points*-Methode bestimmt werden. Über die Eigenschaft Threshold kann der Schwellenwert für die in dieser Arbeit beschriebene Schnittpunktmethode mit der Amplifikationsfunktion eingestellt werden. Über die Eigenschaft AutoAdjust wird festgelegt, ob der Schwellenwert so gesetzt wird, dass der Fehler der Standardgeraden minimiert wird.
TQuantData	Verwaltet die zur Quantifizierung benötigten Daten einer Probe. Führt die Hintergrundkorrektur der Fluoreszenzdaten durch, bestimmt die exponentielle Phase, berechnet Amplifikationsfunktion, Effizienz, C_T-Wert, Konzentration und das Signalrauschen. Über die nur-lesen-Eigenschaften Error und ResultStatus lassen sich Fehlerwerte und Status der Quantifizierung abfragen.
TQuantGroup	Berechnet die statistischen Kennwerte zu den C_T-Werten und Konzentrationen einer Probengruppe. Die Eigenschaft SamplesValid zeigt an, ob sich in der Gruppe eine oder mehrere Proben befinden, die nicht ordnungsgemäß analysiert werden konnten.
TRun	Verwaltung der LightCycler-Daten.
TRunInfo	Enthält allgemeine Informationen zum Experiment.
TSample	Verwaltet die Daten einer Probe. Über die Eigenschaft RawData vom Typ TSegDataList ist der Zugriff auf alle Messdaten dieser Probe möglich.
TSegData	Speichert die Daten eines Segments eines LightCycler-Programms.
TSegDataList	Verwaltet eine Liste von Objekten des Typs TSegData.
TSegList	Speichert eine Liste von Objekten des Typs TSegment.
TSegment	Speichert die Protokolldaten eines Segments eines LightCycler-

Klassenbezeichner	Beschreibung
	Programms.
TSmpList	Speichert eine Liste von Objekten des Typs TSample.

Tab. 7-3 Felder der Klassen zur Verwaltung von LightCycler-Daten.

Feldbezeichner	Werte	Bedeutung
TRun		
Filename	Zeichenkette	Dateiname der geladenen LightCycler-Datei
CCSet[1]	Matrix von Fließkommawerten	Parametersatz zur Farbkompensation
TRunInfo		
User	Zeichenkette	Benutzer des LightCyclers (z. B. „User1“)
Date	Datumswert	Datum des Experiments
Time	Zeitwert	Uhrzeit des Experiments
Units[1]	Zeichenkette	Einheiten zur Konzentrations-/Mengenbestimmung bei der Quantifizierung (z. B. „Kopien“)
CalLED	Fließkommawert	Kalibrierwert der LED
CalGains	Liste von Fließkommawerten	Vorverstärkungsfaktoren der Photohybriden
AutoGain	Wahrheitswert	automatische Einstellung der Verstärkungsfaktoren ein/aus
ContMove	Wahrheitswert	Kontinuierliche Rotation des Probenkarussells ein/aus
Machine	Ganzzahl	Kennziffer des LightCyclers
Notes[1]	Zeichenkette	Notizen zum Experiment
Version	Zeichenkette	Versionskennung der LighCycler Run Software
CCSets[1]	Liste mit Matrizen von Fließkommawerten	Liste der Parameter der importierten Farbkalibrierungsdaten
TRunProgram		
Name	Zeichenkette	Bezeichnung des Programms (z.B. „Ampli“)
CycleCount	Ganzzahl	Anzahl der Zyklen des Programms
AnMode	TAnMode[2]	Analyse-Modus des Programms

Feldbezeichner	Werte	Bedeutung
Channel	Ganzzahl	Kennziffer für den Kanal bzw. das Kanalverhältnis, der bzw. das ausgewertet werden soll
TSegment		
Temp1	Fließkommazahl	Zieltemperatur in °C
Temp2	Fließkommazahl	Zweite Zieltemperatur in °C
StepSize	Fließkommazahl	Temperaturinkrement zum Erreichen der zweiten Zieltemperatur in K
StepDelay	Fließkommazahl	Anzahl Zyklen Verzögerung bis zum ersten Temperaturinkrement
Duration	Fließkommazahl	Dauer in s, über welche die Zieltemperatur gehalten wird
Ramp	Fließkommazahl	Temperaturgradient zum Einstellen der Zieltemperatur
AcMode	TAcMode[3]	Acquisitionsmodus
DataIndex	Ganzzahl	Index der Segmentdaten in der Liste RawData der Probe
DataCount	Ganzzahl	Anzahl der aufgezeichneten Messwerte
TSample		
Name	Zeichenkette	Bezeichnung der Probe (z.B. „Sample1“)
Comment[1]	Zeichenkette	Kommentar zur Probe
ID	Ganzzahl	Position im Probenkarussell
SampleType	TSampleType[4]	Probentyp
ReplicaOf	Ganzzahl	ID der Probe, zu deren Gruppe sie gehört
StdConc	Fließkommazahl	Konzentration des Standards
UserData[1,5]	Zeiger	Benutzerdefinierte Daten

[1] Optionale Angabe.

[2] Selbstdefinierter Datentyp: [ANNONE, ANREGULAR, ANQUANTIFICATION, ANMELTING]

[3] Selbstdefinierter Datentyp: [ACNONE, ACSINGLE, ACSTEP, ACCONTINOUS]

[4] Selbstdefinierter Datentyp: [STSTANDARD, STUNKNOWN, STPOSITIVE, STNEGATIVE, STCONTROL]

[5] Die Daten sind nicht Bestandteil der LightCycler-Daten. Sie werden im Programm z. B. dazu verwendet, zu vermerken, ob die Probe zur Auswertung verwendet wird.

Tab. 7-4 Felder der Klassen zur Analyse und Auswertung von LightCycler-Daten.

Feldbezeichner	Werte	Bedeutung
TAnalysis		
DataSourceX	TSegDataType[1]	Typ der Segmentdaten, die als x-Werte verwendet werden
DataSourceY	TSegDataType[1]	Typ der Segmentdaten, die als y-Werte verwendet werden
ChannelIndex	16bit-Ganzzahl	Kennziffer der auszuwertenden Kanäle[2]
TAnalysisData		
Status	TAnalysisDataStatus[3]	Status der Auswertung
TQuantData		
AmpData	Liste von Fließkommawerten	Amplifikationsdaten
IndExp	Paar von Ganzzahlenwerten	Indexwert für Zyklus mit maximaler Krümmung der Amplifikationskurve
ExpRange	Paar von Fließkommawerten	Bereich exponentiellen Anstiegs der Amplifikationskurve
Efficiency	Fließkommazahl	Effizienz anhand der Steigung der exponentiellen Phase der Amplifikationskurve
AmpFunc	TFunc[4]	Funktion zur Beschreibung der Amplifikationskurve im Bereich ihres exponentiellen Anstiegs
CT	Fließkommazahl	C_T-Wert
Conc	Fließkommazahl	Berechnete Konzentration
CV	Fließkommazahl	Variationskoeffizient der Konzentration
Noise	Fließkommazahl	Rauschgrenze
Error	TQuantDataError[5]	Fehlerindikator
ResultStatus	TQuantDataResultStatus[6]	Statusindikator des Ergebnisses
TMeltData		
PeakData	TFltList	Datenreihe der Peakdaten (2. Ableitung der Schmelzkurvendaten)
Peaks	TPeaks[7]	Funktion zur Beschreibung der Peaks (Summe von Gaußfunktionen)

[1] Selbstdefinierter Datentyp: [sdtCycles, sdtTime, sdtTemp, sdtChannels]

[2] Das höherwertige Byte gibt die Nummer des Numerator-Kanals an, das niederwertige die des Denominator-Kanals zur Bildung von Verhältnissen aus den Fluoreszenzdaten zweier Kanäle (Null = kein Kanal).

[3] Selbstdefinierter Datentyp: [adsOK, adsInvalidData]

[4] Selbstdefinierte Klasse zur Verwaltung von Funktionen

[5] Selbstdefinierter Datentyp: [qdeNoError, qdeNoCT, qdeNoExp]

[6] Selbstdefinierter Datentyp: [qrsPositive, qrsNegative, qrsLowConc, qrsHighConc, qrsInvalidThreshold, qrsInvalidNoiseband, qrsLowEfficiency, qrsHighEfficiency, qrsInvalid]

[7] Selbstdefinierte Klasse zur Verwaltung der Summe von Gaußfunktionen: Siehe 0.

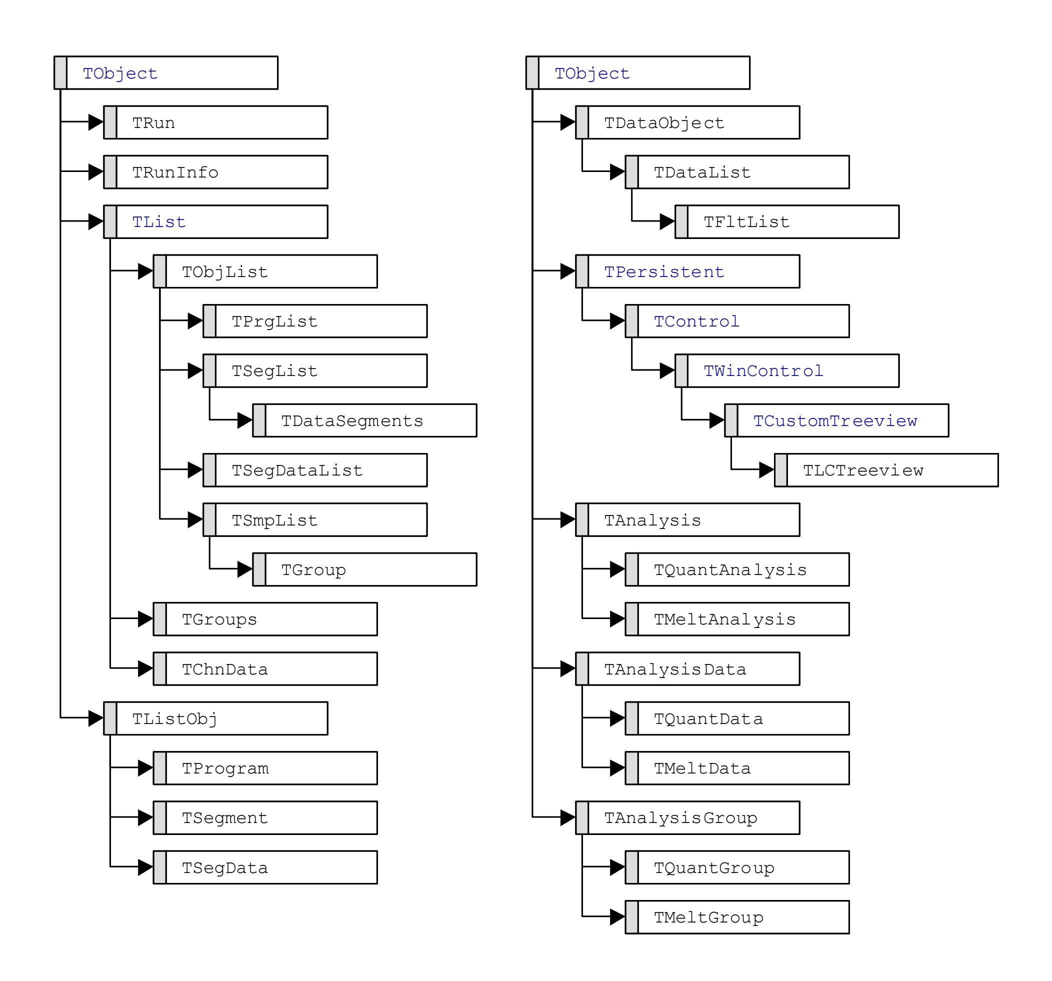


Abb. 7-1 Stammbaum der Klassen zur Verwaltung und Auswertung von LightCycler-Daten. Die blau beschrifteten Klassen sind in Delphi vordefinierte Klassen. Die auf der linken Seite aufgeführten Klassen werden zur Verwaltung der LightCycler-Daten verwendet. Auf der rechten Seite ist die Ableitungsfolge für TFltList, TLCTreeview und die zur Analyse der Daten benötigten Klassen gezeigt. Die Klasse TFltList verwaltet eine Liste von Fließkommawerten variabler Größe. Sie wird überall verwendet, wo Datenreihen zur Analyse oder zur Darstellung zu verwalten sind. Die Klasse TLCTreeView verwaltet die geöffneten LightCycler-Dateien sowie die Selektion der Proben. Die von TAnalysis, TAnalysisData und TAnalysisGroup abgeleiteten Klassen verwalten die analysebezogenen Daten zur Quantifizierung bzw. Schmelzkurvenanalyse und führen die nötigen Berechnungen durch.

Abschließend sei noch auf einen wichtigen Zusammenhang hingewiesen, der als Dilberts „Salary Theorem“ bekannt ist:

Dilbert's "Salary Theorem" states that "engineers and scientists can never earn as much as business executives and sales people." This theorem can now be supported by a mathematical equation based on the following two postulates:

Postulate 1: Knowledge is Power.

Postulate 2: Time is Money.

As every engineer knows:

Power = Work / Time.

Since

Knowledge = Power

and

Time = Money,

it is obvious, that

Knowledge = Work/Money.

Solving for Money, we get:

Money = Work / Knowledge.

Thus, as Knowledge approaches zero, Money approaches infinity, regardless of the amount of work done.

Conclusion: The less you know, the more you make.

[Author unknown]